

MODELOWANIE I PROGNOZOWANIE

MATERIAŁY DYDAKTYCZNE

- o Podstawowe charakterystyki danych statystycznych,
- o Indeksy,
- o Prognozowanie- wiadomości wstępne,
- o Modele ekonometryczne,
- o Jednorównaniowy model liniowy,
- o Prognoza na podstawie modelu liniowego,
- o Model tendencji rozwojowej,
- o Prognozowanie na podstawie szeregów czasowych,
- o Naiwne i proste metody prognozowania,
- o Model Browna,
- o Model Holta,
- o Analiza wariantowa,
- o Teoria gier,
- o Zestawy danych statystycznych.

Lucjan Kowalski

Warszawa 2010

PODSTAWOWE CHARAKTERYSTYKI DANYCH STATYSTYCZNYCH

RODZAJE ŚREDNICH

Jedną z wielkości charakteryzujących dane liczbowe jest wartość średnia.

Rodzaje podstawowych średnich:

- arytmetyczna
- geometryczna
- harmoniczna

Wybór średniej zależy od rodzaju badanych wielkości i potrzeb analizy danych.

Najczęściej stosowaną średnią jest średnia arytmetyczna.

Średnią arytmetyczną liczb rzeczywistych $x_1, x_2, x_3, \dots, x_n$ nazywamy liczbę:

$$\bar{x} = \frac{1}{n} (x_1 + x_2 + \dots + x_n) = \frac{1}{n} \sum_{i=1}^n x_i$$

Przykład.

W pięciu wybranych stolicach liczba linii metra wynosi: 3, 2, 5, 2, 3. Ile wynosi średnia linii metra w tych stolicach?

(odp. 3)

Jeżeli wśród n danych występują wartości powtarzające się:

$$x_i \text{ występuje } n_i \text{ razy, } i = 1, 2, \dots, r \qquad n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i = n$$

to

$$\bar{x} = \frac{1}{n} (n_1 x_1 + n_2 x_2 + \dots + n_k x_k) = \frac{1}{n} \sum_{i=1}^k n_i x_i = \sum_{i=1}^k \frac{n_i}{n} x_i$$

Ten sposób liczenia średniej arytmetycznej nazywamy **średnią arytmetyczną ważoną**.

Przykład.

W dwudziestu pięciu wybranych państwach liczba instytutów medycznych badających nowy wirus grypy jest następująca:

w dziesięciu po 3 instytuty,

w dziesięciu po 4 instytuty,

w pięciu po 6 instytutów.

Ile wynosi średnia liczba tych instytutów w rozpatrywanych państwach?

$$\bar{x} = \frac{1}{25} (10 \cdot 3 + 10 \cdot 4 + 5 \cdot 6) = \frac{100}{25} = 4$$

Średnią geometryczną liczb rzeczywistych dodatnich $x_1, x_2, x_3, \dots, x_n$ nazywamy pierwiastek n – tego stopnia z ich iloczynu, tzn.

$$\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} = \sqrt[n]{\prod_{i=1}^n x_i}$$

Średnia geometryczna znajduje najczęściej zastosowanie przeciętnego tempa zmian w czasie, np. do uśredniania indeksów łańcuchowych.

Przykład.

Roczny procentowy przyrost liczby turystów odwiedzających dany region w czterech kolejnych latach wynosił: 10%, 20%, 5%, 15%. Jaki był średni przyrost w tym okresie?

$$\bar{x}_g = \sqrt[4]{1,1 \cdot 1,2 \cdot 1,05 \cdot 1,15} = \sqrt[4]{1,5939} \approx 1,1236$$

Średni procentowy przyrost liczby turystów w tym okresie wynosił około 12,36% z roku na rok.

Zauważmy, że średnia arytmetyczna tych danych wynosi 12,5%

Jeżeli wśród n danych występują wartości powtarzające się:

$$x_i \text{ występuje } n_i \text{ razy, } i = 1, 2, \dots, r \quad n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i = n$$

to

$$\bar{x}_g = \sqrt[n]{(x_1)^{n_1} \cdot (x_2)^{n_2} \cdot \dots \cdot (x_k)^{n_k}} = \sqrt[n]{\prod_{i=1}^k (x_i)^{n_i}}$$

Ten sposób liczenia średniej geometrycznej nazywamy **średnią geometryczną ważoną**.

Średnią harmoniczną liczb $x_1, x_2, x_3, \dots, x_n$ różnych od zera nazywamy odwrotność średniej arytmetycznej odwrotności liczb, tzn.

$$\bar{x}_h = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Średnią harmoniczną stosuje się przy uśrednianiu wielkości względnych, np. przy obliczaniu przeciętnej prędkości lub średniej gęstości zaludnienia.

Przykład.

Gęstość zaludnienia w trzech 200-tysięcznych miastach wynosi odpowiednio 100, 300 i 600 osób km^2 . Obliczmy przeciętną gęstość zaludnienia.

$$H = \frac{3}{\frac{1}{100} + \frac{1}{300} + \frac{1}{600}} = \frac{3}{\frac{6+2+1}{600}} = \frac{1800}{9} = 200 \text{ osób/km}^2$$

Stosując średnią arytmetyczną otrzymalibyśmy:

$$\bar{x} = \frac{100 + 300 + 600}{3} = \frac{1000}{3} = 333 \text{ osób/km}^2$$

Przykład.

Pan Kowalski codziennie dojeżdża do pracy samochodem z prędkością 40km/h. Pewnego dnia zasnął i wyjechał później niż zwykle. W połowie trasy zorientował się, że nie zdąży i zwiększył prędkość o 20km/h, dzięki czemu nie spóźnił się do pracy. Z jaką średnią prędkością jechał tego dnia pan Kowalski?

$$\bar{x}_h = \frac{2}{\frac{1}{40} + \frac{1}{60}} = \frac{240}{5} = 48$$

Zauważmy, że średnia arytmetyczna tych danych wynosi 50km/h

Jeżeli wśród n danych występują wartości powtarzające się:

$$x_i \text{ występuje } n_i \text{ razy, } i = 1, 2, \dots, r \quad n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i = n$$

to

$$\bar{x}_h = \frac{n}{\frac{n_1}{x_1} + \frac{n_2}{x_2} + \dots + \frac{n_k}{x_k}} = \frac{n}{\sum_{i=1}^k \frac{n_i}{x_i}}$$

Ten sposób liczenia średniej harmonicznego nazywamy **średnią harmoniczną ważoną**.

Twierdzenie

Dla dowolnych liczb rzeczywistych dodatnich $x_1, x_2, x_3, \dots, x_n$ zachodzą nierówności

$$\bar{x}_h \leq \bar{x}_g \leq \bar{x}$$

przy czym równość zachodzi wtedy i tylko wtedy, gdy $x_1 = x_2 = x_3 = \dots = x_n$.

Zróznicowanie danych.

Zróznicowanie danych (rozrzut) mierzymy uśredniając ich odchylenie od średniej.

Wariancja

$$s^2 = \frac{1}{n} [(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Np. dla danych 2, 5, 3, 4, 6, średnia wynosi 4. Aby wyznaczyć wariancję liczymy sumę kwadratów odchyleń poszczególnych danych od średniej:

$$(2-4)^2 + (5-4)^2 + (3-4)^2 + (4-4)^2 + (6-4)^2 = 4 + 1 + 1 + 0 + 4 = 10$$

otrzymana sumę dzielimy przez 5 (liczba danych). Zatem wariancja dla powyższych danych wynosi 2.

Jeżeli wśród n danych występują wartości powtarzające się:

$$x_i \text{ występuje } n_i \text{ razy, } i = 1, 2, \dots, r \qquad n_1 + n_2 + \dots + n_k = \sum_{i=1}^k n_i = n$$

to

$$s^2 = \frac{1}{n} (n_1(x_1 - \bar{x})^2 + n_2(x_2 - \bar{x})^2 + \dots + n_k(x_k - \bar{x})^2) = \frac{1}{n} \sum_{i=1}^k n_i(x_i - \bar{x})^2$$

Uwaga

$$s^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2 = \frac{1}{n} \sum_{i=1}^k n_i x_i^2 - (\bar{x})^2$$

Wariancja mierzy rozrzut (zróźnicowanie) danych statystycznych (punktem odniesienia jest średnia) lecz miara ta wyrażona jest w kwadratach jednostek rozpatrywanych danych statystycznych co utrudnia interpretację, dlatego w praktyce częściej stosujemy pierwiastek z wariancji nazywany odchyleniem standardowym.

Odchylenie standardowe

$$s = \sqrt{s^2},$$

Współczynnik zmienności

$$v = \frac{s}{\bar{x}} \qquad (\text{niekiedy wynik jest podawany w procentach})$$

Współczynnik zmienności mierzy zróźnicowanie względne i określa jaką część (ile procent) przeciętnego poziomu badanej cechy stanowi odchylenie standardowe.

Ponieważ jest to wielkość niemianowana, często bywa stosowany do porównywania zróźnicowania dwóch cech lub tej samej cechy w różnych populacjach.

Przedział typowych wartości

$$[\bar{x} - s, \bar{x} + s],$$

Jest to przedział do którego należy większość danych statystycznych, interpretacja ta jest uzasadniona wtedy gdy cecha ma rozkład zbliżony do rozkładu normalnego.

Rozstęp

$$r_0 = x_{\max} - x_{\min},$$

SZEREG CZASOWY

Wartości zjawiska tworzą szereg czasowy (szereg dynamiczny):

t_i	t_1	t_2		t_n
y_i	y_1	y_2		y_n

t_i – chwile lub okresy (przedziały powinny być jednakowe)

Uwaga: niekiedy stosuje się zapis: $t_0, t_1, \dots, t_n$.

INDEKSY (wskaźniki dynamiki)

Indeksy mierzą zmianę poziomu zjawiska między dwoma wyróżnionymi okresami (momentami).

Indeksy dzielimy na:

- indeksy indywidualne (proste),
- indeksy zespołowe (agregatywne).

INDEKSY INDYWIDUALNE

Indeksy indywidualne stosujemy przy badaniu dynamiki zjawisk jednorodnych.

a) **ciąg indeksów o stałej podstawie:**

$$I_{1/0}, I_{2/0}, I_{3/0}, \dots, I_{n/0}$$

y_0 – stała podstawa (dowolna spośród $y_1, \dots, y_n$).

gdzie $I_{t/0} = \frac{y_t}{y_0} \quad (t = 1, 2, \dots, n)$

(można dodać $I_{0/0} = 1 = 100\%$)

b) **ciąg indeksów łańcuchowych:**

$$I_{2/1}, I_{3/2}, I_{4/3}, \dots, I_{n/n-1}$$

gdzie

$$I_{t/t-1} = \frac{y_t}{y_{t-1}} \quad (t = 2, 3, \dots, n)$$

Przykład

Y – liczba wypadków drogowych w ciągu roku.

Rok	t	liczba wypadków y_t	$I_{t/0} = \frac{y_t}{y_0}$ $y_0 = 2005$	$I_{t/t-1} = \frac{y_t}{y_{t-1}}$
2005	1	39 779	1,000	—
2006	2	40 373	1,015	1,015
2007	3	43 755	1,100	1,084
2008	4	38 832	0,976	0,887
2009	5	40 454	1,017	1,042

Średnie tempo dynamiki to średnie tempo zmian przypadające na jednostkę czasu.

Zagadnienie

Wyznaczyć liczb $\bar{g}$ taką, że gdyby wszystkie indeksy łańcuchowe były sobie równe i miały wartość $\bar{g}$ to startując z wartości y_1 obliczona wartość zjawiska w okresie t_n byłaby równa y_n (taka sama jak przy różnych indeksach łańcuchowych).

Liczbę $\bar{g}$ nazywamy **średnim tempem dynamiki lub średnim tempem zmian lub średnim indeksem łańcuchowym**.

Zauważmy, że

$$(*) \quad y_n = I_{n/n-1} \cdot \dots \cdot I_{2/1} \cdot y_1$$

gdyby wszystkie indeksy były równe to

$$(**) \quad y_n = \bar{g}^{n-1} \cdot y_1$$

Porównując (*) i (**) mamy

$$\bar{g} = \sqrt[n-1]{I_{n/n-1} \cdot \dots \cdot I_{2/1}}$$

(średnia geometryczna)

Własność:

$$\bar{g} = \sqrt[n-1]{I_{n/1}} = \sqrt[n-1]{\frac{y_n}{y_1}}$$

Średni wskaźnik tempa to

$$\bar{T} = \bar{g} - 1$$

Dodatni wskaźnik tempa oznacza, że średnio zjawisko wzrastało.

Ujemny wskaźnik tempa oznacza, że średnio zjawisko malało.

Przykład

Dla danych z poprzedniego przykładu:

$$\bar{g} = \sqrt[4]{1,015 \cdot 1,084 \cdot 0,887 \cdot 1,042} = \sqrt{\sqrt{1,017}} = 1,004 = 100,4\%$$

$$\bar{T} = 0,004 = 0,4\%$$

Oznacza to, że średnio z roku na rok liczba wypadków wzrastała o 0,4%.

Uwaga

Średnie tempo dynamiki można stosować do wyznaczania wartości zjawiska w okresach następnych (ostatnią wartość zjawiska mnożymy przez odpowiednią potęgę $\bar{g}$).

Przykład

Dla danych z poprzedniego przykładu wyznacz prognozę liczby wypadków drogowych w roku 2010 i 2011.

Prognozowanie- wiadomości wstępne

Prognozowanie to racjonalne wnioskowanie o zdarzeniach nieznanych na podstawie zdarzeń znanych.

Celem prognozy jest dostarczenie obiektywnych informacji potrzebnych do podejmowania decyzji.

Prognozy a symulacje.

Prognoza – *co będzie w momencie t*,

Symulacja – *co by było gdyby*

Przykład

Z rozpatrywanego modelu wynika, że wydatki na prasę i książki stanowią 5% miesięcznych dochodów rodziny. Ustalono, że miesięczne dochody rodziny wyniosą 4000 zł. Możemy zatem postawić **prognozę**, że wydatki na prasę i książki wyniosą 200 zł.

Jeśli jednak wyznaczalibyśmy wydatki na prasę i książki dla różnych wariantów dochodu, np.

wydatki 190 zł dla dochodu 3800,

wydatki 200 zł dla dochodu 4000,

wydatki 220 zł dla dochodu 4400,

to byłyby **symulacje**.

Procedury prognozowania

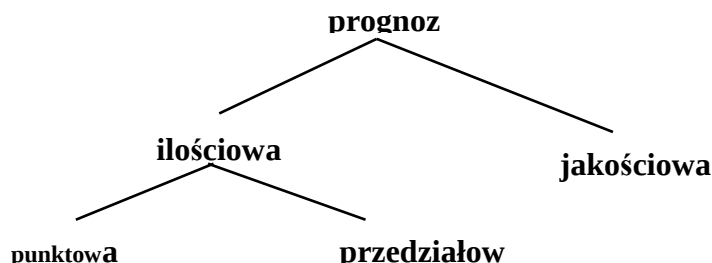
- Proste i intuicyjne (na podstawie prostych charakterystyk liczbowych),
- Ekonometryczne,
- Poprzez analogię,
- Prognozy ekspertów (heurystyczne),
- Wyznaczanie różnych scenariuszy rozwoju.

Prognozowane zmiany wartości badanego zjawiska mogą być:

- ilościowe (zgodne z dotychczasową prawidłowością np. trendem lub funkcją regresji),
- jakościowe (odejście od dotychczasowych prawidłowości)

Uproszczona klasyfikacja prognoz.

Ze względu na wartości prognozy:



Ze względu na okres prognozy:

- **Krótkookresowa** (na taki okres w którym mogą zachodzić tylko zmiany ilościowe),
- **Średniookresowa** (na taki okres w którym mogą zachodzić zmiany ilościowe i niewielkie zmiany jakościowe),
- **Długookresowa** (na taki okres w którym mogą zachodzić zarówno zmiany ilościowe jak i jakościowe).

W praktyce niekiedy podział ten odnosi się do zasięgu ekstrapolacji (liczba jednostek czasu wyjścia z prognoza w przyszłość) w porównaniu z liczbą danych:

do 10% - prognoza **krótkookresowa**,

od 10 do 20% - prognoza **średniookresowa**,

powyżej 20% - prognoza **długookresowa**,

Ponieważ wartości prognoz wyznaczamy w oparciu o dane, to muszą być one dobrej jakości.

Cechy danych decydujące o ich jakości:

- rzetelność,
- jednoznaczność,
- identyfikowalność,
- kompletność,
- aktualność,
- koszt (zbierania i opracowania),
- porównywalność (np. w zakresie: czasowym, terytorialnym, pojęciowym).

Etapy prognozowania:

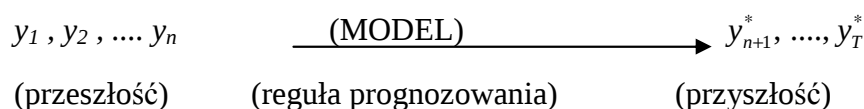
- Sformułowanie zadania prognostycznego
 - Określenie zmiennych prognozowanych,
 - Ustalenie celu prognozy,
 - Ustalenie horyzontu prognozy i warunków jej dopuszczalności
- Określenie przesłanek prognostycznych
 - Określenie czynników kształtujących badane zjawisko,
 - Zbieranie danych,
- Wybór metody prognozowania
- Wyznaczanie prognoz
- Ocena dopuszczalności prognoz
- Wykorzystanie prognozy
- Weryfikacja i monitorowanie (przy powtarzalności) prognozy.

Podstawowy schemat prognozowania.

Y - badane zjawisko,

y_t - obserwacje badanego zjawiska,

y_t^* - prognozowane wartości badanego zjawiska.



Bezwzględny błąd prognozy jest równy $y_\tau^* - y_\tau$,

Względny błąd prognozy jest równy $\frac{|y_\tau^* - y_\tau|}{y_\tau}$ (ma zwykle sens dla zjawisk o wartościach

dodatnich), można go wyrażać w procentach.

gdzie y_τ to prawdziwa wartość zjawiska w okresie prognozy.

Uwaga

Bezwzględny błąd prognozy niekiedy definiuje się jako $y_\tau - y_\tau^*$.

Względny błąd prognozy niekiedy definiuje się jako $\frac{|y_\tau^* - y_\tau|}{y_\tau^*}$.

Prawdziwą wartość błędu prognozy można wyznaczyć dopiero po ustaleniu prawdziwej wartości badanego zjawiska, wcześniej błąd można tylko **oszacować**.

Szacowanie błędu prognozy.

1. Na podstawie prognoz wygasłych (ex post),
2. Metoda stochastyczna (ex ante).

Ad. 1. Wykorzystuje się informacje o trafności prognozowania w przeszłości. Przyjmuje się, że trafność prognoz przyszłych będzie podobna do trafności prognoz przeszłych. Prognozy wygasłe używane do szacowania powinny być wyznaczone w ten sam sposób jak ostateczna prognoza. Jako oszacowanie błędu prognozy można np. przyjąć średnią z

modułów błędów bezwzględnych $\frac{1}{k} \sum_{t=1}^k |y_t^* - y_t| * 100\%$ lub względnych

$\frac{1}{k} \sum_{t=1}^k \left| \frac{y_t^* - y_t}{y_t} \right| * 100\%$ prognoz wygasłych. Ten sposób szacowania błędu prognozy

zastosujemy przy modelach adaptacyjnych.

Ad. 2. Wykorzystuje się stochastyczne założenia o stosowanym modelu. Przyjmuje się, że błąd prognozy jest zbliżony do średniej rozbieżności między możliwymi wartościami prognozowanego zjawiska a możliwymi prognozami tego zjawiska w okresie prognozy.

Jako oszacowanie błędu prognozy można np. błąd średniokwadratowy $\sqrt{\frac{1}{k} \sum_{t=1}^k (y_t^* - y_t)^2}$

lub względny błąd średniokwadratowy $\sqrt{\frac{1}{k} \sum_{t=1}^k \left(\frac{y_t^* - y_t}{y_t} \right)^2} * 100\%$. Ten sposób szacowania

błędu prognozy zastosujemy przy modelach ekonometrycznych.

Niekiedy przyjmuje się, że prognoza jest **dopuszczalna**, gdy szacowany błąd nie przekracza 5 – 10%.

Schemat prognozowania na podstawie modelu ekonometrycznego $y = f(x)$

x_t^* - wektor zmiennych objaśniających dla okresu prognozy.

Prognoza punktowa: $y_t^* = f(x_t^*)$.

Prognoza przedziałowa: $\langle y_t^* - \Delta_1, y_t^* + \Delta_2 \rangle$,

Zwykle $\Delta_1 = \Delta_2 = \Delta$ (= błąd bezwzględny prognozy przedziałowej).

Jakość prognozy w znacznym stopniu zależy od jakości zastosowanego modelu ekonometrycznego.

Oprócz tego

- Błąd prognozy powinien być mały,
- Przyjęte wartości zmiennych objaśniających powinny być wiarygodne,
- Okres prognozy powinien być sensowny.

Przykład

Rozpatrując model $Y = 200 - 10X$,

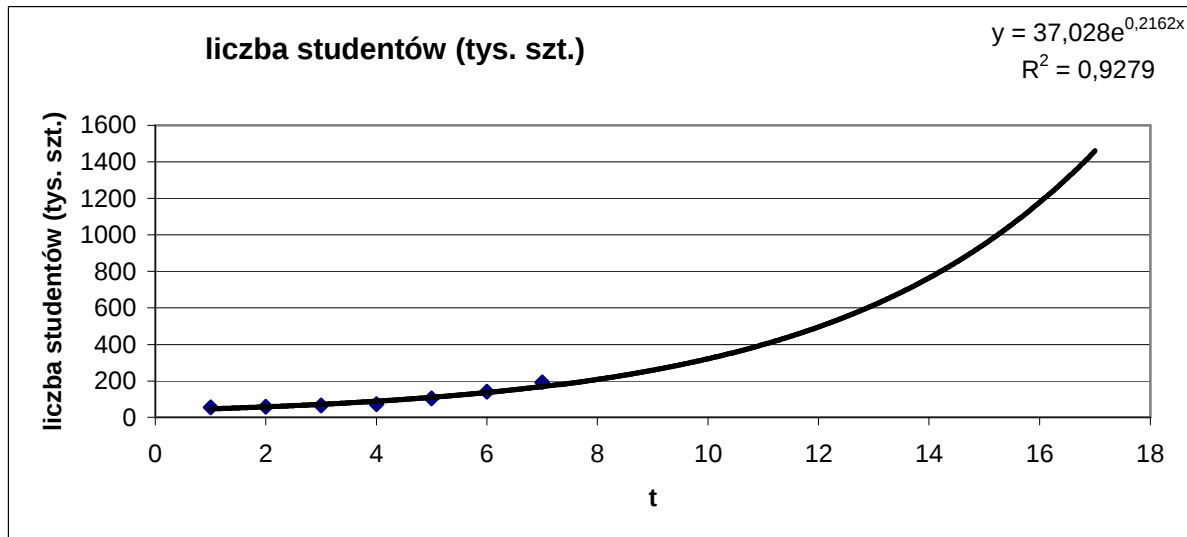
Y – jednostkowe koszty produkcji,

X – wielkość produkcji

W tym przypadku prognozy tracą sens dla $x > 20$.

Przykład

Liczba studentów kierunków ekonomicznych w Polsce (tys. osób) liczona na koniec roku akademickiego w latach 1991-97 wynosiła: 54, 58, 65, 71, 104, 140, 193.



Rozpatrując model $Y = f(t)$,

Y – liczba studentów,

t – rok

W tym przypadku prognoza np. na rok 2007 (ponad 1,46 mln osób) byłaby przesadna.

MODELE EKONOMETRYCZNE

Model ekonometryczny to opis stochastycznej zależności badanego zjawiska ekonomicznego od czynników kształtujących go, wyrażony w postaci równości lub układu równości.

Jeśli np. rozpatrujemy zjawisko popytu na określony towar lub grupę towarów i przyjmujemy, że głównym czynnikiem kształtującym popyt jest cena to możemy rozpatrywać model

$$D = f(P) \quad D - \text{popyt}, \quad P - \text{cena.}$$

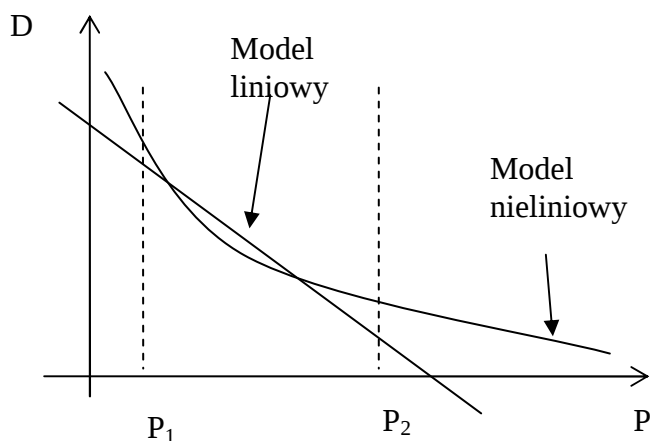
Z prawa malejącego popytu wynika, że funkcja f powinna być malejąca ($(P_1 < P_2 \Rightarrow f(P_1) > f(P_2))$). Zależność tę możemy zrealizować za pomocą różnych funkcji malejących, najprostsza z nich to funkcja liniowa:

$$D = a + bP \quad (\text{liniowy model popytu}), \quad a > 0; \quad b < 0$$

jeśli model liniowy nie pasuje do zaobserwowanych wielkości to należy zastosować model nieliniowy np. model potęgowej:

$$D = a \cdot P^b \quad (\text{potęgowy model popytu}), \quad a > 0; \quad b < 0$$

Dla pewnych zakresów cen model liniowy może być dobrym przybliżeniem modelu nieliniowego



Niekiedy model z jedną zmienną źle opisuje badane zjawisko, wtedy możemy rozpatrywać model z wieloma zmiennymi. W modelu popytu drugim czynnikiem kształtującym popyt może być dochód, wtedy rozpatrujemy zależność:

$$D = f(P, I) \quad I - \text{dochód ludności.}$$

Zależność tę możemy jak poprzednio zrealizować za pomocą funkcji liniowej

$$D = a + bP + cI$$

lub potęgowej

$$D = a \cdot P^b I^c$$

Ogólna postać modelu w postaci jednej równości:

$$Y = f(X, \varepsilon)$$

X, Y - zmienne, (X może być postaci $X = (X_1, X_2, \dots, X_k)$), ε - element losowy

gdy

$$Y = f(X) + \varepsilon \text{ to } \varepsilon \text{ nazywamy składnikiem losowym,}$$

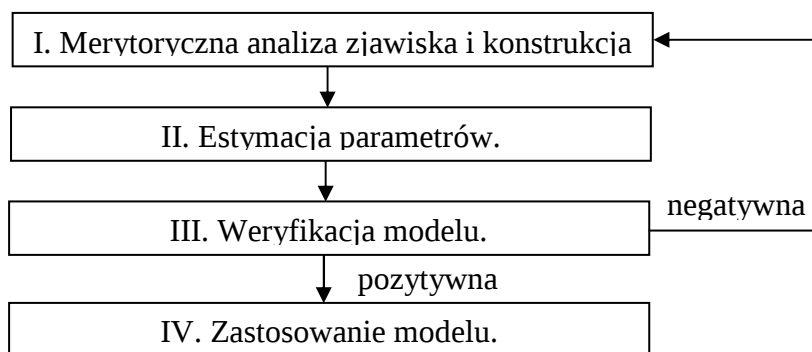
gdy

$$Y = f(X)\varepsilon \text{ to } \varepsilon \text{ nazywamy czynnikiem losowym.}$$

Powody uwzględniania elementu losowego w modelu ekonometrycznym:

- nie uwzględnienie wszystkich czynników kształtujących badane zjawisko (najczęściej nie uwzględniamy czynników mających mały wpływ i element losowy reprezentuje łączny wpływ takich zmiennych),
- możliwość występowania błędów w pomiarze wielkości zmiennych,
- brak pewności czy przyjęta do obliczeń postać funkcyjna modelu jest prawidłowa.

Etapy modelowania ekonometrycznego:



Uproszczona **klasyfikacja zmiennych** w modelu

- zmienna endogeniczna – zmienna, której wartości określone są w modelu,
- zmienna egzogeniczna – zmienna, której wartości określone są poza modelem,
- zmienna objaśniana – występuje po lewej stronie równań modelu,
- zmienna objaśniająca – występuje po prawej stronie równań modelu.

Każda ze zmiennych może być bieżąca lub opóźniona.

Uwaga:

W modelach wielowymiarowych zmienna objaśniana może być jednocześnie zmienną objaśniającą.

Przykład 1.

Rozpatrzmy model wzrostu gospodarczego

$$\begin{cases} DN_t = aNI_{t-4}^b Z_t^c \varepsilon_{1t} \\ NI_t = dDN_t + \varepsilon_{2t} \end{cases}$$

gdzie

DN - dochód narodowy,

NI - nakłady inwestycyjne,

Z - zatrudnienie,

a, b, c, d - parametry strukturalne,

$\varepsilon_1, \varepsilon_2$ - elementy losowe (ε_1 - czynnik losowy, ε_2 - składnik losowy)

Klasyfikacja:

- zmienne endogeniczne: DN_t, NI_t, NI_{t-4}
- zmienne egzogeniczne: Z_t
- zmienne objaśniane: DN_t, NI_t
- zmienne objaśniające: NI_{t-4}, Z_t, DN_t
- zmienne bieżące: DN_t, Z_t, NI_t
- zmienne opóźnione: NI_{t-4} .

Klasyfikacja modeli

Modele klasyfikujemy ze względu na następujące kryteria:

a) liczba zależności w modelu

- modele jednorównaniowe,
- modele wielorównaniowe,

b) postać zależności funkcyjnej,

- modele liniowe,
- modele nieliniowe (potęgowe, wykładnicze, itp.).

c) rola czasu w równaniach,

- modele statyczne (nie uwzględniają czasu),
- modele dynamiczne.

Przykład 2

Model z przykładu 1 jest:

- dwurównaniowy,
- nieliniowy,
- dynamiczny.

Przykład 3 (model popytu)

$$D = a + bP + cI + \varepsilon$$

D - popyt,

P - cena,

I - dochód ludności.

Jest to model:

- jednorównaniowy,
- liniowy,
- statyczny.

Przykład 4 (model równowagi rynkowej)

$$\begin{cases} D = a + bP + cI + \varepsilon_1 \\ S = d + eP + \varepsilon_2 \\ D = S \end{cases}$$

S - sprzedaż.

Przykład 5 (Model wielkości produkcji Cobba-Douglasa)

$$P_t = aX_t^b Y_t^{1-b} \quad a > 0, \quad 0 < b < 1$$

P - wielkość (wartość) produkcji przedsiębiorstwa,

X - zatrudnienie (fundusz płac),

Y - wartość środków trwałych

Jednorównaniowy model liniowy z jedną zmienną objaśniającą

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

gdzie:

Y - zmienna objaśniana, y_i - wartości (obserwacje) zmiennej Y; $i = 1, \dots, n$ - numer obserwacji,

X - zmienna objaśniająca, x_i - wartości zmiennej X,

β_0, β_1 - parametry strukturalne (ich przybliżoną wartość wyznacza się na podstawie obserwacji (x_i, y_i))

ε - składnik losowy.

Zakładamy, że

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, 2, \dots, n$$

tzn. każda zaobserwowana wartość y_i jest funkcją liniową x_i z dokładnością do składnika losowego ε_i .

Zakładamy również, że x_i są ustalonymi wartościami (nielosowymi), takimi samymi w powtarzalnych próbach. Składniki losowe ε_i są losowymi zmiennymi niezależnymi o zerowej wartości przeciętnej i wariancji, która nie zależy od i (homoskedastyczność).

Aby wyznaczyć przybliżoną wartość parametrów strukturalnych β_0, β_1 na podstawie próby stosujemy metodę najmniejszych kwadratów (MNK).

MNK polega na wyznaczeniu takich przybliżeń

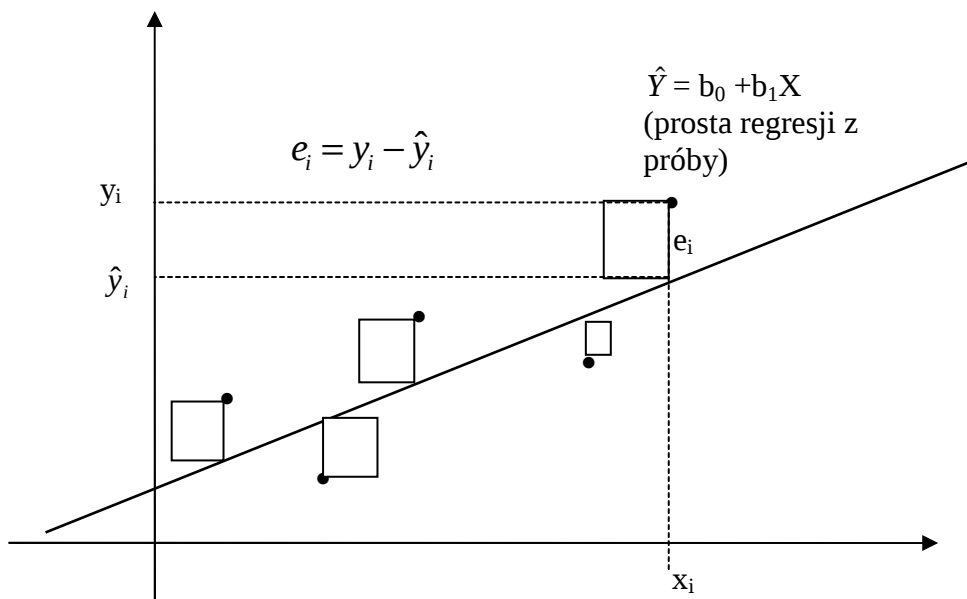
$$b_0 \approx \beta_0 \quad b_1 \approx \beta_1$$

aby dla danych obserwacji (x_i, y_i) suma kwadratów odchyłeń zaobserwowanych wartości y_i od wartości teoretycznych $\hat{y}_i = \beta_0 + \beta_1 x_i$ była minimalna, tzn. chcemy wyznaczyć minimum funkcji:

$$(*) \quad S(b_0, b_1) = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - b_0 - b_1 x_i)^2$$

$e_i = y_i - \hat{y}_i$ nazywamy **resztami** modelu regresji

MNK:



Należy wyznaczyć prostą regresji tak aby suma pól kwadratów była minimalna.

Obliczając pochodne cząstkowe funkcji (*) i przyrównując do zera otrzymujemy (układ równań normalnych)

$$\frac{\partial S}{\partial \beta_0} = 2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)(-1) = -2 \left(\sum_{i=1}^n y_i - \beta_1 \sum_{i=1}^n x_i - n\beta_0 \right) = 0$$

$$\frac{\partial S}{\partial \beta_1} = 2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)(-x_i) = -2 \left(\sum_{i=1}^n y_i x_i - \beta_1 \sum_{i=1}^n x_i^2 - \beta_0 \sum_{i=1}^n x_i \right) = 0$$

rozwiązując otrzymany układ równań (np. pierwsze równanie mnożymy przez $-\bar{x} = -\frac{1}{n} \sum_{i=1}^n x_i$ i dodając stronami obliczamy β_1) otrzymamy wzory na przybliżone wartości parametrów strukturalnych

$$b_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} = \frac{\frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum x_i^2 - (\bar{x})^2} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Można wykazać, że dla tych wartości spełniony jest warunek dostateczny minimum.
Prostą

$$\hat{Y} = b_0 + b_1 X$$

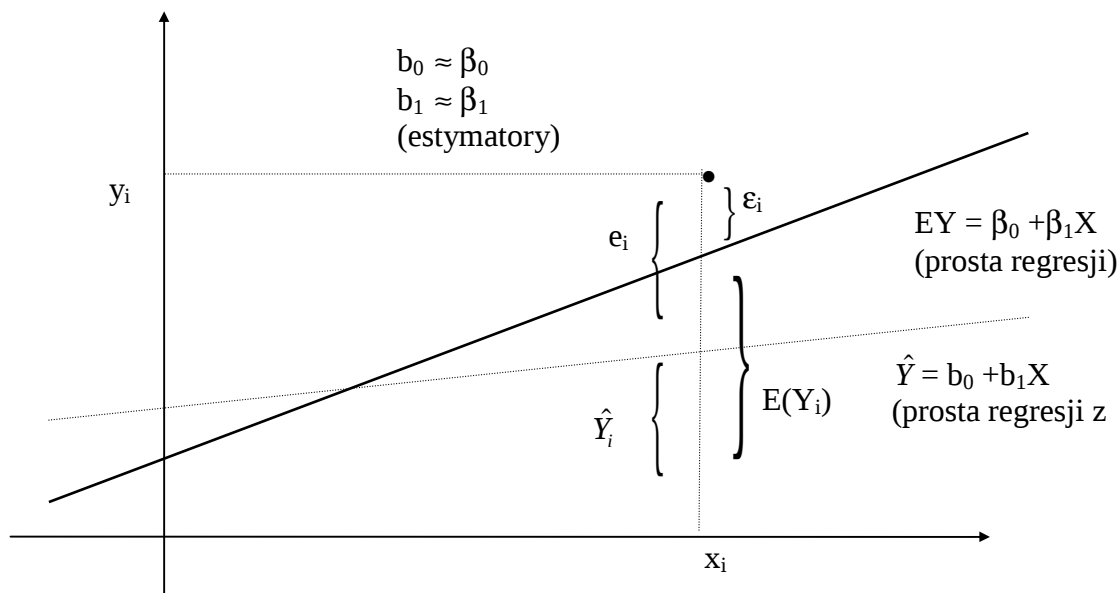
nazywamy **prostą regresji z próby**.

Uwaga:

$$a) \quad \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - n\bar{x}\bar{y}$$

$$b) \quad \sum (x_i - \bar{x})^2 = \sum x_i^2 - n(\bar{x})^2$$

Model regresji liniowej:



Uwaga

Gdy X jest zmienną czasową $x_i = t$ tzn. model ma postać $Y = \beta_0 + \beta_1 t + \varepsilon$ wówczas taki model nazywamy **modelem tendencji rozwojowej** lub **modelem trendu liniowego**.

Wtedy korzystając z ustalonych wartości t wzory na b_1 i b_0 można uprościć (patrz odpowiedni temat).

Miary dopasowania.

Wariancja resztowa:

Wariancja resztowa to uśrednienie pól kwadratów zbudowanych na resztach i odzwierciedla stopień dopasowania prostej regresji do danych statystycznych.

Niech, $e_i = y_i - \hat{y}_i$, gdzie $\hat{y}_i = b_0 + b_1 x_i$ wtedy

$$S_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$$

czyli

$$S_e^2 = \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{n-2}$$

$S_e = \sqrt{S_e^2}$ oznacza średnie (standardowe) odchylenie od prostej regresji.

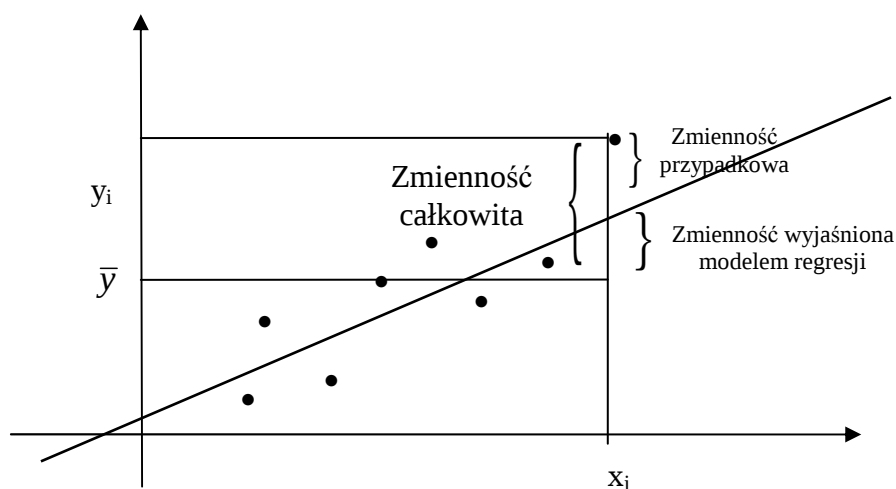
Dopasowanie modelu do danych empirycznych można oceniać odchyleniem standardowym reszt lecz jest to miara bezwzględna i nieunormowana, dlatego do porównań lepsze są miary względne lub unormowane.

Najprostszą względną miarą dopasowania jest **współczynnik zmienności losowej** :

$$V_e = \frac{S_e}{\bar{Y}} 100\%$$

Współczynnik ten informuje jaką część średniej wartości badanego zjawiska stanowi odchylenie standardowe reszt.

Mniejsze wartości tego współczynnika wskazują na lepsze dopasowanie modelu do danych empirycznych, niekiedy żąda się aby np. $V_e < 0,2$ (0,3).



Wprowadzamy oznaczenia:

Całkowita suma kwadratów (zmienność całkowita): $CSK = \sum (y_i - \bar{y})^2$

Wyjaśniona suma kwadratów (zmienność wyjaśniona): $WSK = \sum (\hat{y}_i - \bar{y})^2$

Niewyjaśniona suma kwadratów (zmienność przypadkowa): $NSK = \sum e_i^2$

gdzie : $\hat{y}_t = b_0 + b_1 x_t$

Własność: $\sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum e_i^2$

Czyli $CSK = WSK + NSK$

Miarą dopasowania modelu do rzeczywistości (wartości zaobserwowanych) jest również współczynnik determinacji R^2

Współczynnik determinacji: $R^2 = \frac{WSK}{CSK}$ $R^2 \in \langle 0, 1 \rangle$

Współczynnik ten określa jaką część całkowitej zmienności zmiennej objaśnianej została wyjaśniona przez model regresji liniowej.

Praktyczne sposoby obliczania współczynnika determinacji:

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} = \frac{b_0 \sum y_i + b_1 \sum x_i y_i - n \bar{y}^2}{\sum y_i^2 - n(\bar{y})^2} =$$

$$= \frac{b_1 (\sum x_i y_i - n \bar{x} \bar{y})}{\sum y_i^2 - n(\bar{y})^2} = 1 - \frac{S_e^2}{\frac{n}{n-2} S_Y^2} = \frac{\text{cov}^2(X, Y)}{S_X^2 S_Y^2} = r^2$$

Wniosek. Dla modelu liniowego wartość liczbową współczynnika determinacji jest równa kwadratowi współczynnika korelacji Pearsona.

Przykład

Badano zależności kosztów całkowitych (w tys. zł.) Y od wielkości produkcji (tys. szt.) X w 6-ciu zakładach produkcyjnych.

x_t	2	4	3	2	6	1
y_t	2	5	4	4	7	2

Dla modelu $Y = \beta_0 + \beta_1 x + \varepsilon$ wyznaczamy przybliżone wartości parametrów strukturalnych i współczynnik determinacji.

Obliczenia wykonamy w tabeli

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$\hat{y}_i$	$\hat{y}_i - \bar{y}$	$(\hat{y}_i - \bar{y})^2$
2	2	-1	-2	2	1	4	3	-1	1
4	5	1	1	1	1	1	5	1	1
3	4	0	0	0	0	0	4	0	0
2	4	-1	0	0	1	0	3	-1	1
6	7	3	3	9	9	9	7	3	9
1	2	2	-2	4	4	4	2	-2	4
18	24	0	0	16	16	18	24	0	16

$$\bar{x} = \frac{18}{6} = 3; \quad \bar{y} = \frac{24}{6} = 4; \quad b_1 = \frac{16}{16} = 1; \quad b_0 = 4 - 1 \cdot 3 = 1$$

zatem związek pomiędzy kosztami całkowitymi a wielkością produkcji wyraża się zależnością liniową w postaci

$$\hat{Y} = 1 + X$$

Współczynnik determinacji

$$R^2 = \frac{16}{18} = 0,89$$

należy oczekiwać, że rozpatrywany model wyjaśnia 89% całkowitej zmienności kosztów całkowitych produkcji.

Prognoza na podstawie modelu liniowego. (jedna zmienna)

$\hat{Y} = b_0 + b_1 X$ oszacowany model ekonometryczny.

Prognoza punktowa.

Niech X_τ – przewidywana wartość cechy X w okresie prognozy.

Prognoza punktowa y_τ^* to przewidywana wartość cechy Y odpowiadająca wartości X_τ cechy X .

$$y_\tau^* = b_0 + b_1 x_\tau$$

Standardowy błąd prognozy

$$s_\tau = s_e \sqrt{1 + \frac{1}{n} + \frac{(x_\tau - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} = s_e \sqrt{1 + \frac{\sum_{i=1}^n x_i^2 + nx_\tau^2 - 2x_\tau \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}}$$

Zatem należy traktować wartość prognozy jako $y_\tau^* \pm s_\tau$

Gdzie $s_e = \sqrt{s_e^2}$ to odchylenie resztowe.

Niech $e_i = y_i - \hat{y}_i$, gdzie $\hat{y}_i = b_0 + b_1 x_i$ wtedy

$$s_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2}$$

czyli

$$s_e^2 = \frac{\sum_{i=1}^n y_i^2 - b_0 \sum_{i=1}^n y_i - b_1 \sum_{i=1}^n x_i y_i}{n-2}$$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{s_\tau}{y_\tau^*} \cdot 100\%$$

Model tendencji rozwojowej

Gdy X jest zmienną czasową $x_i = t$ ($t = 1, 2, \dots, n$) tzn. model regresji ma postać $\hat{Y} = b_0 + b_1 t$ wówczas taki model nazywamy **modelem tendencji rozwojowej** lub **modelem trendu liniowego**

Wtedy korzystając z własności:

$$(*) \quad \sum_{t=1}^n t = \frac{n(n+1)}{2}, \quad \sum_{t=1}^n t^2 = \frac{n(n+1)(2n+1)}{6}, \quad \bar{t} = \frac{n+1}{2}$$

$$\sum (t - \bar{t})^2 = \sum t^2 - n(\bar{t})^2$$

Mamy

$$b_1 = \frac{n \sum t y_t - \sum t \sum y_t}{n \sum t^2 - (\sum t)^2} = \frac{12 \sum (t - \bar{t}) y_t}{n(n^2 - 1)}$$

$$b_0 = \bar{y} - b_1 \bar{t} = \bar{y} - b_1 \frac{n+1}{2}$$

Wariancja resztowa

Niech $e_i = y_i - \hat{y}_i$, (gdzie $\hat{y}_i = b_0 + b_1 x_i$) to reszty modelu, wtedy

$$s_e^2 = \frac{\sum_{i=1}^n e_i^2}{n-2} \quad \text{czyli} \quad s_e^2 = \frac{\sum_{t=1}^n y_t^2 - b_0 \sum_{t=1}^n y_t - b_1 \sum_{t=1}^n t y_t}{n-2}$$

$S_e = \sqrt{s_e^2}$ oznacza średnie (standardowe) odchylenie od trendu liniowego.

Dopasowanie modelu do danych empirycznych oceniamy też współczynnikiem determinacji

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2} = 1 - \frac{\sum e_i^2}{\sum (y_i - \bar{y})^2} =$$

$$= \frac{b_0 \sum y_i + b_1 \sum t y_t - n \bar{y}^2}{\sum y_t^2 - n(\bar{y})^2} = \frac{b_1 (\sum t y_t - n \bar{t} \bar{y})}{\sum y_t^2 - n(\bar{y})^2} = r^2$$

Prognoza dla modelu trendu

Niech t_τ – okres prognozy.

Prognoza punktowa y_τ^* to przewidywana wartość cechy Y w okresie t_τ .

$$y_\tau^* = b_0 + b_1 t_\tau$$

Standardowy błąd prognozy punktowej

$$s_{\tau} = s_e \sqrt{1 + \frac{1}{n} + \frac{(t_{\tau} - \bar{t})^2}{\sum_{t=1}^n (t - \bar{t})^2}} = s_e \sqrt{1 + \frac{\sum_{t=1}^n t^2 + nt_{\tau}^2 - 2t_{\tau} \sum_{t=1}^n t}{n \sum_{t=1}^n t^2 - \left(\sum_{t=1}^n t \right)^2}}$$

Wzór ten można uprościć korzystając z własności (*).

$$s_{\tau} = s_e \sqrt{1 + \frac{\frac{n(n+1)(2n+1)}{6} + nt_{\tau}^2 - 2t_{\tau} \frac{n(n+1)}{2}}{\frac{n^2(n+1)(2n+1)}{6} - \frac{n^2(n+1)^2}{4}}} =$$

$$= s_e \sqrt{1 + \frac{2(2n+1) + \frac{12t_{\tau}^2}{n+1} - 12t_{\tau}}{n^2 - n}}$$

Zatem należy traktować wartość prognozy jako $y_{\tau}^* \pm s_{\tau}$

Jakość prognozy punktowej możemy ocenić względnym błędem prognozy punktowej

$$\delta_{punkt} = \frac{s_{\tau}}{y_{\tau}^*} \cdot 100\%$$

Przykład.

Liczba maturzystów (tys. osób) w pewnym województwie w latach 1991-1999 wynosiła;

Rok	1991	1992	1993	1994	1995	1996	1997	1998	1999
Y	3,8	6	7,5	7,5	9,2	9,9	22,3	13,2	13,4

Otrzymane równanie trendu liniowego ma postać

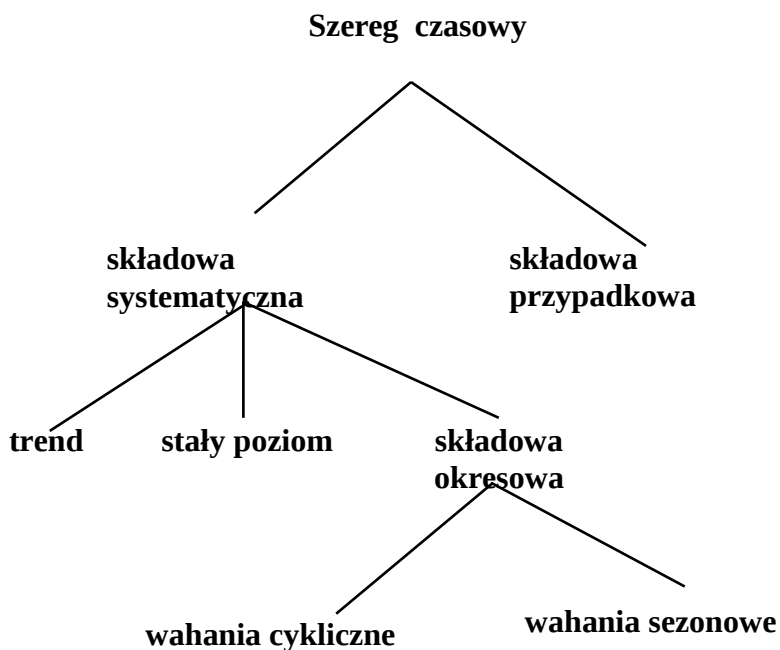
$$\hat{Y} = 3,2 + 1,2t$$

Przewidywana liczba maturzystów w latach następnych:

Rok 2000 - 15200 osób, rok 2001 - 164000 osób.

Prognozowanie na podstawie szeregów czasowych.

Składowe szeregi czasowych.

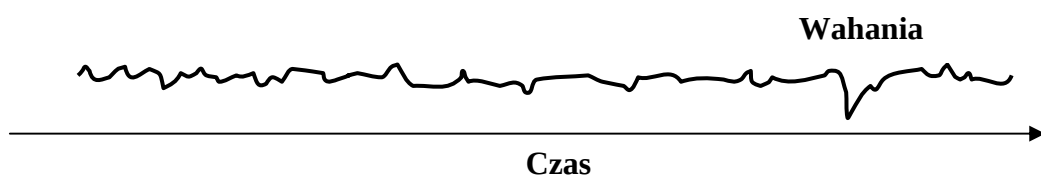
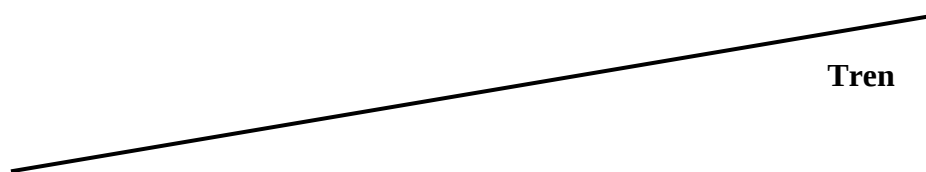
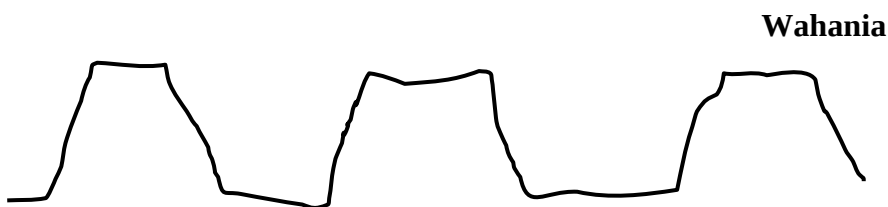
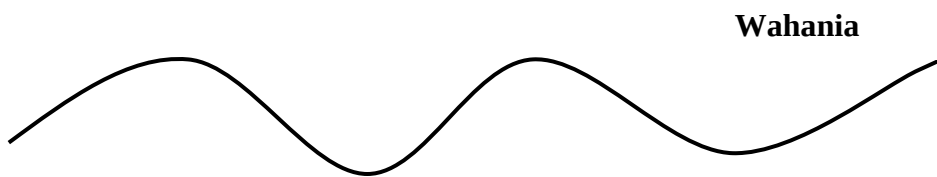
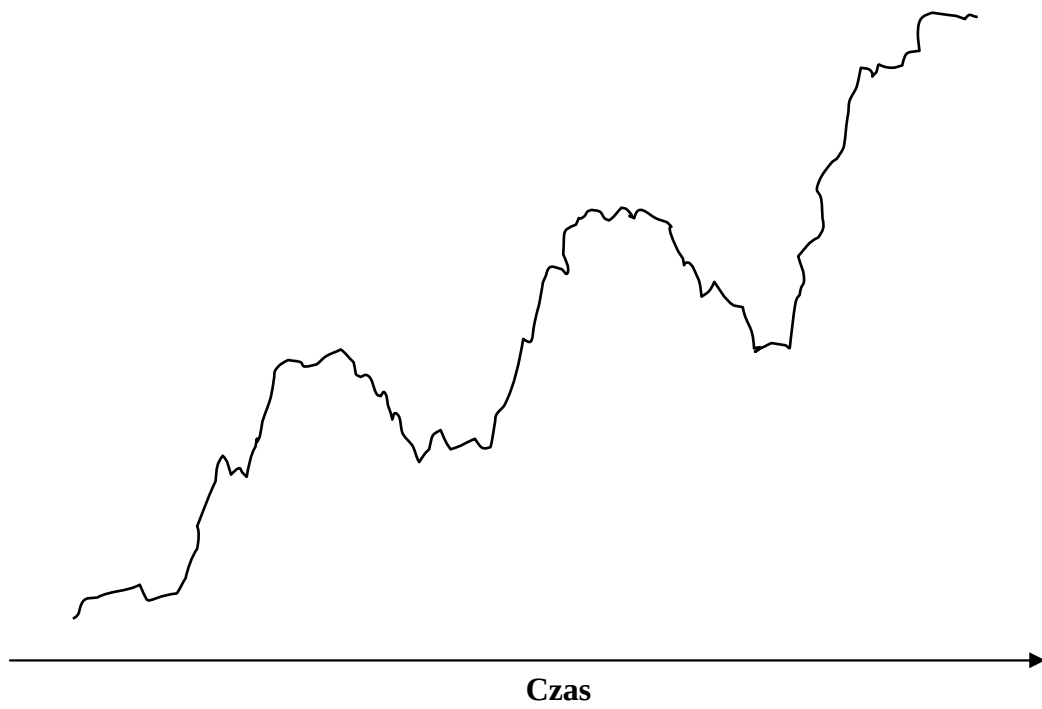


Trend (tendencja rozwojowa) - długookresowa skłonność do jednokierunkowych zmian (wzrostu lub spadku) wartości zmiennej badanej. Jest konsekwencją działania stałych czynników np. w przypadku sprzedaży - liczba potencjalnych klientów, ich dochody lub preferencje. Może być wyznaczony gdy mamy długi ciąg obserwacji.

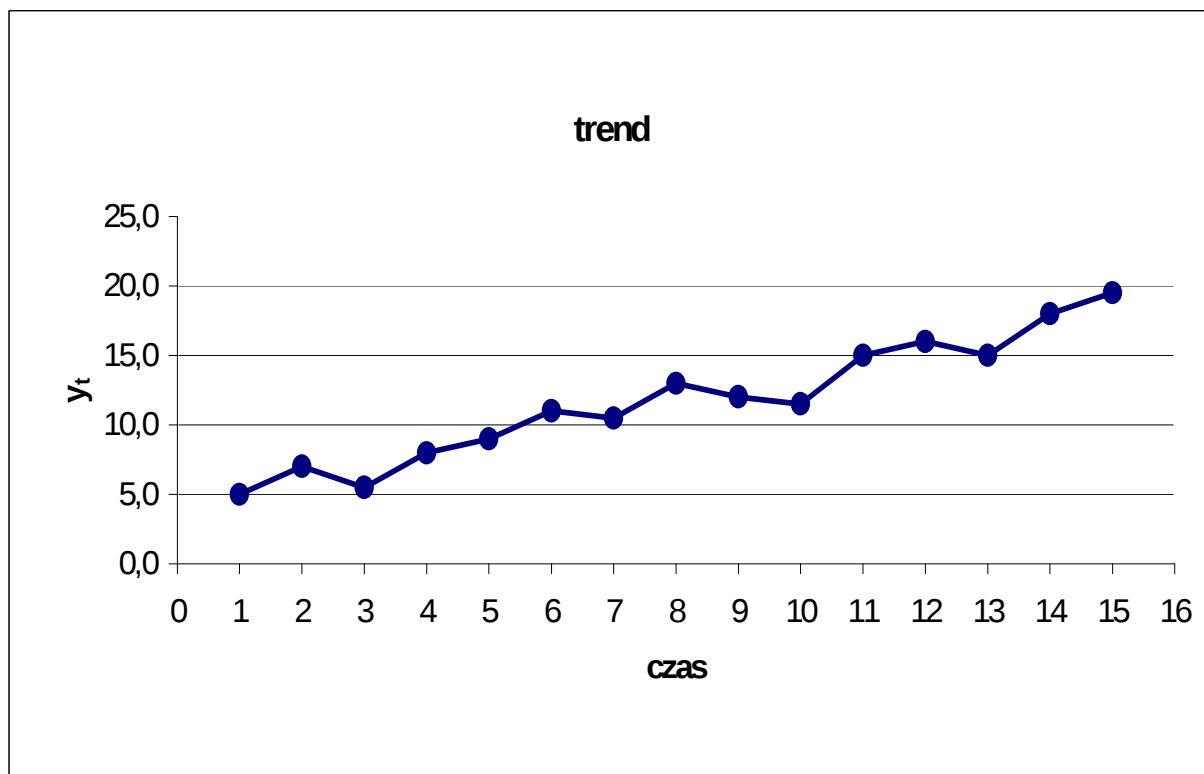
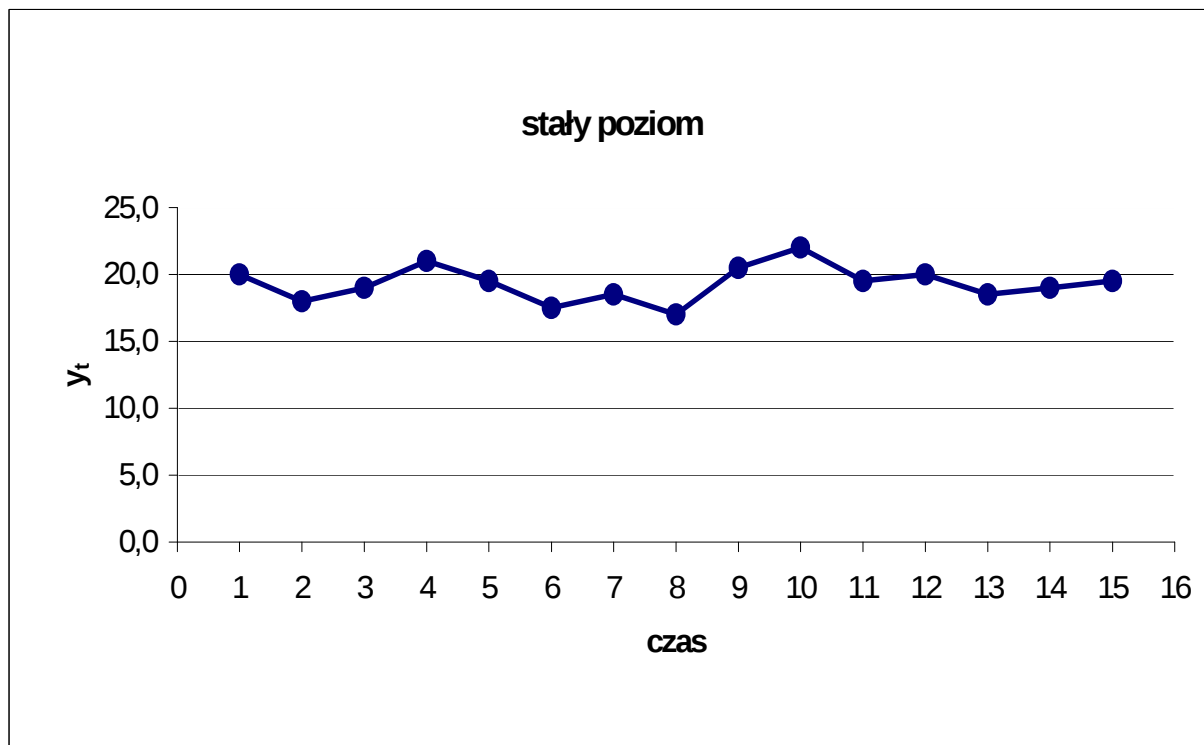
Stały (przeciętny poziom) - występuje gdy w szeregu czasowym nie ma trendu, zaś wartości badanej zmiennej oscylują wokół pewnego stałego poziomu.

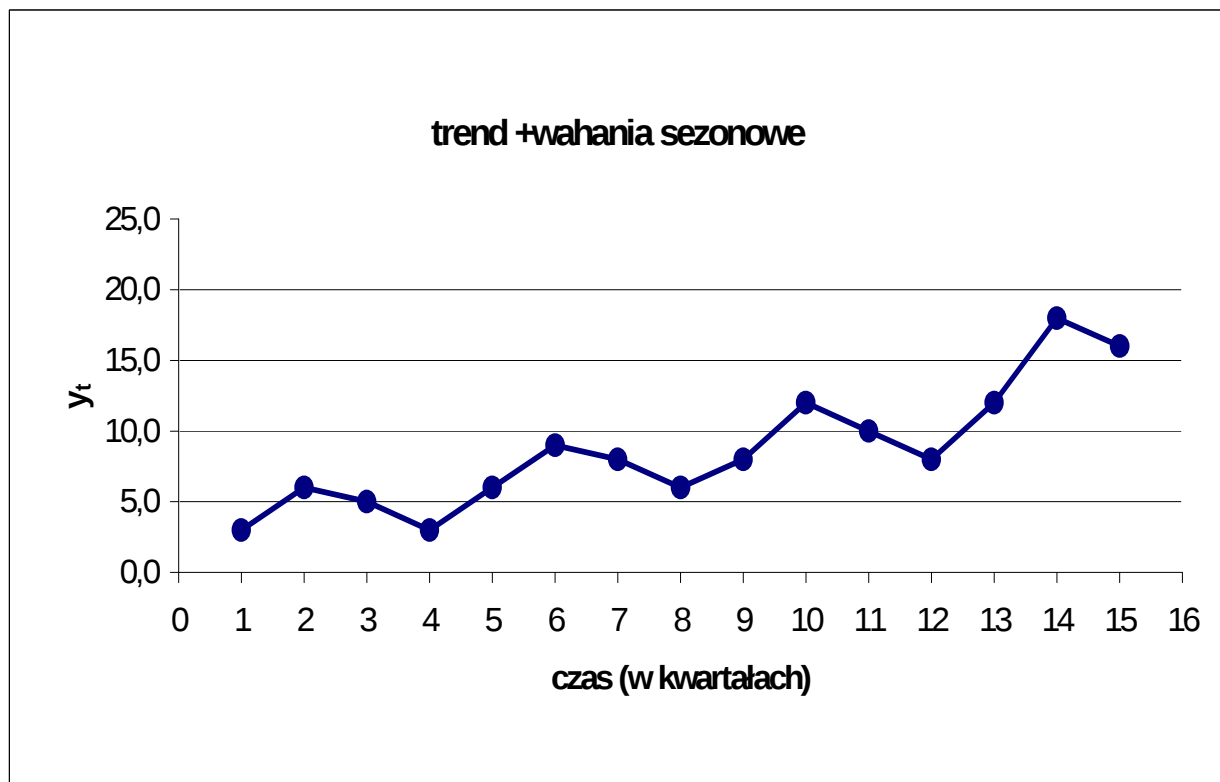
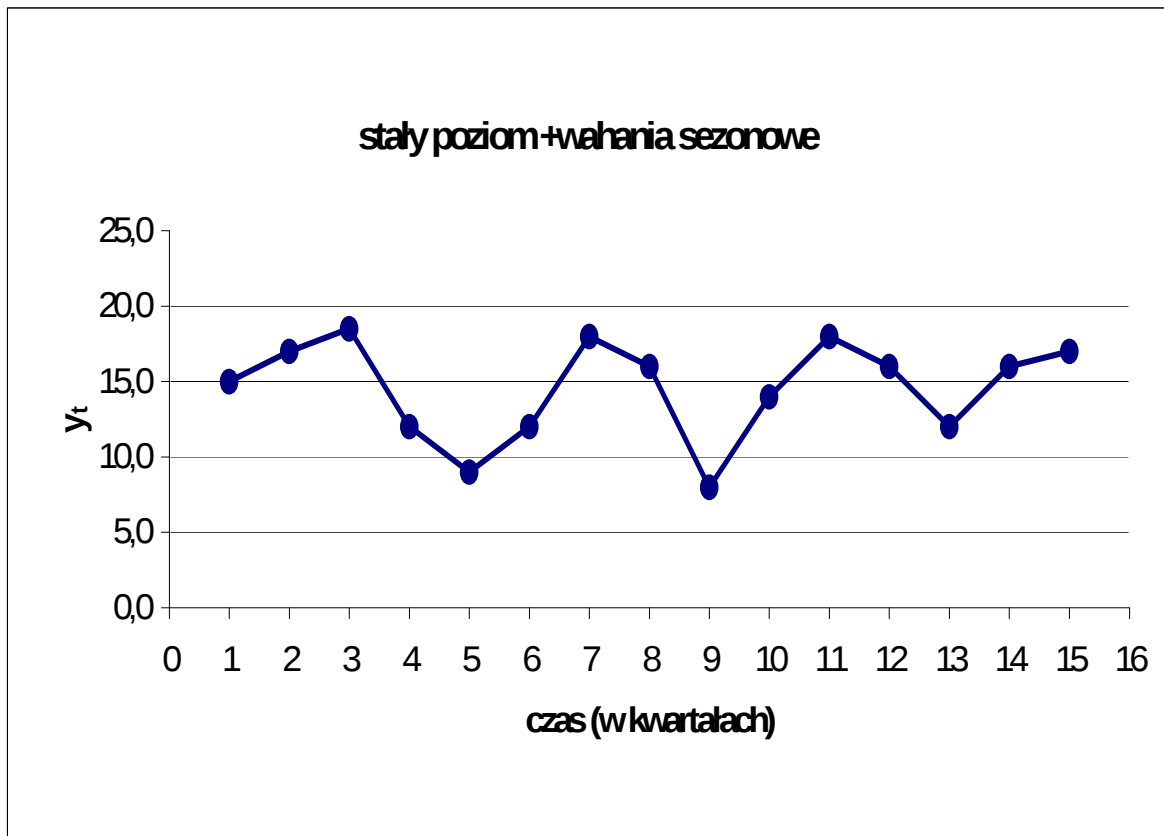
Wahania cykliczne - długookresowe wahania wokół trendu lub stałego poziomu. W ekonomii najczęściej związane z cyklem koniunkturalnym gospodarki.

Wahania sezonowe - wahania wokół trendu lub stałego poziomu. Wahania te mają skłonności do powtarzania się w określonym czasie nie przekraczającym jednego roku, odzwierciedlają wpływ pogody lub kalendarza na działalność gospodarczą.



Dekompozycja szeregu czasowego = wyodrębnienie poszczególnych składowych.





Niekiedy przyjmuje się, że wahania przypadkowe są niewielkie, gdy ich współczynnik zmienności jest rzędu kilku, najwyżej kilkunastu procent.

Prognoza zmiennej Y jest wartością funkcji f zależnej od czasu, przeszłych wartości i prognoz tej zmiennej.

$$y_t^* = f(t, \underbrace{y_{t-1}, \dots, y_{t-p}}_{\text{wartości}}, \underbrace{y_{t-1}^*, \dots, y_{t-p}^*}_{\text{prognozy}}, \underbrace{\xi_t}_{\text{składnik losowy}})$$

p - wielkość opóźnienia.

Uwaga.

Jakość modelu oceniamy jak w ekonometrii.

Dopasowanie wielkości zjawiska wyznaczonych z modelu (prognozy wygasłe) do wielkości zaobserwowanych oceniamy na podstawie:

- **błędu średniokwadratowego** prognoz wygasłych

$$s^* = \sqrt{\frac{1}{k} \sum_{t=1}^k (y_t - y_t^*)^2} \quad k - \text{liczba prognoz wygasłych}$$

wielkość ta określa o ile średnio jednostek prognozy wygasłe odchylają się (plus-minus) od wartości zaobserwowanych.

- **względego błędu średniokwadratowego** (procentowego błędu średniokwadratowego) prognoz wygasłych

$$s_w^* = \sqrt{\frac{1}{k} \sum_{t=1}^k \left(\frac{y_t - y_t^*}{y_t} \right)^2} \cdot 100\%$$

wielkość ta określa o ile średnio procent prognozy wygasłe odchylają się (plus-minus) od wartości zaobserwowanych.

- **średniego błędu względnego** (procentowego błędu względnego) prognoz wygasłych

$$\psi = \frac{1}{k} \sum_{t=1}^k \left| \frac{y_t - y_t^*}{y_t} \right| \cdot 100\%$$

(interpretacja jak wyżej)

Często ostatni z tych błędów służy do oceny jakości prognozy.

Naiwne i proste metody prognozowania.

(oparte na założeniu, że wahania przypadkowe są niewielkie i nie zmienia się dotychczasowy wpływ czynników kształtujących obserwowane zjawisko).

Zaletą metody naiwnej jest prostota, wadą brak oceny jakości prognozy na podstawie prognoz wygasłych.

Rodzaje prognoz naiwnych:

- wg stałego poziomu

$$y_{t+1}^* = y_t \quad \text{lub} \quad y_{t+1}^* = \frac{y_t + y_{t-1} + y_{t-2}}{3}$$

- wg stałych przyrostów bezwzględnych (np. trend zbliżony do liniowego)

$$y_{t+1}^* = y_t + (y_t - y_{t-1})$$

- wg stałych przyrostów względnych (niektóre trendy nieliniowe)

$$y_{t+1}^* = y_t \left(1 + \frac{y_t - y_{t-1}}{y_{t-1}} \right) = \frac{y_t^2}{y_{t-1}}$$

- wg wahań sezonowych i stałego poziomu

$$y_{t+1}^* = y_{t-r+1}, \quad \text{gdzie } r \text{ długość cyklu sezonowego (liczba faz cyklu),}$$

- wg wahań sezonowych i trendu

$$y_{t+1}^* = y_{t-r+1} + r\Delta\bar{y}_r, \quad \text{gdzie } r \text{ długość cyklu sezonowego (liczba faz cyklu),}$$

$\Delta\bar{y}_r$ - przyrost średnich w dwóch ostatnich cyklach.

Przykład.

Dla poszczególnych serii danych miesięcznych wyznacz prognozę naiwną na kolejny miesiąc.

- 115, 119, 126, 131, 136,
- 1, 4, 12, 24, 72,
- 125, 120, 110, 115, 120.

Przykład.

Wiedząc, że zjawisko ma charakter sezonowy ($r = 4$), dla poszczególnych serii danych kwartalnych wyznacz prognozę naiwną na dwa kolejne kwartały.

- 150, 200, 160, 120, 170,
- 150, 220, 190, 170, 200, 270, 240, 220, 270

Metoda średniej globalnej.

$$y_{n+1}^* = \frac{1}{n} \sum_{i=1}^n y_i$$

Metoda średniej ruchomej.

Metodę tą wykorzystujemy zarówno do wygładzania szeregu czasowego jak i do prognozowania.

Prognoza jest średnią arytmetyczną z k ostatnich obserwacji (k - stała wygładzania).

$$y_t^* = \frac{1}{k} \sum_{i=t-k}^{t-1} y_i$$

k wyznaczamy tak aby średni kwadratowy błąd ex post $s^{*2} = \frac{1}{n-k} \sum_{t=k+1}^n (y_t - y_t^*)^2$ był minimalny.

Prognozę oceniamy za pomocą średniego błędu względnego prognoz przeszłych

$$\Psi_k = \frac{1}{n-k} \sum_{t=k+1}^n \frac{|y_t - y_t^*|}{y_t} 100\%$$

Uwaga.

Gdy k = 1 to metoda naiwna.

Gdy k = n to średnia globalna.

Gdy k duże to średnia ruchoma silniej wygładza szereg czasowy lecz jednocześnie wolniej reaguje na zmiany poziomu badanego zjawiska.

Gdy k małe to średnia ruchoma szybciej odzwierciedla zmiany zjawiska lecz większy wpływ wywierają na nią wahania przypadkowe.

Aby stosować średnią ruchomą powinniśmy zwykle dysponować co najmniej kilkunastoma danymi.

Będziemy stosować zwykle k = 3 lub k = 5.

Średnia ważona.

Ustalamy wagi $0 < w_1 \leq w_2 \leq \dots \leq w_k < 1$ takie, że $\sum_{i=1}^k w_i = 1$ (oznacza to, że do wcześniejszych informacji przywiązujemy mniejszą wagę).

Prognozę wyznaczamy na podstawie wzoru:

$$y_{t+1}^* = \sum_{i=t-k+1}^t w_{i-t+k} y_i$$

Model Browna (prosty model wygładzania wykładniczego).

Zwykle stosujemy ten model dla szeregów czasowych o stałym poziomie lub bardzo słabym trendzie i umiarkowanych wahaniami przypadkowych.

Model pozwala wyznaczyć prognozę wg wzoru:

$$y_t^* = \alpha y_{t-1} + (1-\alpha)y_{t-1}^*, \quad t = 2, 3, \dots, n+1$$

prognoza jest kombinacją wypukłą (średnią ważoną) przeszłej wartości zjawiska i przeszłej prognozy. $\alpha \in \langle 0,1 \rangle$ – parametr wygładzania.

Wartość α dobieramy np. na podstawie kryterium najmniejszego błędu średniokwadratowego prognoz wygasłych s^* tzn. $\min_{\alpha} s^*(\alpha)$ gdzie

$$s^* = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - y_t^*(\alpha))^2}$$

Jeśli nie mamy możliwości wyznaczenia optymalnej wartości parametru wygładzania zwykle zaleca się stosowania wartości 0,1 – 0,3.

Uwaga

Równoważny wzór na prognozę w tym modelu ma postać:

$$y_t^* = y_{t-1}^* + \alpha(y_{t-1} - y_{t-1}^*)$$

zatem dla małych α prognoza w małym stopniu uwzględnia błąd ex post prognoz przeszłych.

Uwaga

Jako wartość y_1^* przyjmujemy jedną z wartości:

a) pierwszą wartość szeregu czasowego, $y_1^* = y_1$,

b) średnią z trzech początkowych wartości szeregu czasowego, $y_1^* = \frac{y_1 + y_2 + y_3}{3}$,

c) średnią z pięciu początkowych wartości szeregu czasowego, $y_1^* = \frac{y_1 + y_2 + y_3 + y_4 + y_5}{5}$.

Model Browna jest rozwinięciem metody średnich ważonych.

Wagi maleją wykładniczo przy coraz starszych danych.

Widać to gdy przekształcimy wzór na prognozę w tym modelu:

$$y_n^* = \alpha y_{n-1} + (1-\alpha)y_{n-1}^*$$

podstawiając $y_{n-1}^* = \alpha y_{n-2} + (1-\alpha)y_{n-2}^*$ otrzymamy

$$y_n^* = \alpha y_{n-1} + (1-\alpha)(\alpha y_{n-2} + (1-\alpha)y_{n-2}^*) = \alpha y_{n-1} + \alpha(1-\alpha)y_{n-2} + (1-\alpha)^2 y_{n-2}^*$$

następnie podstawiając $y_{n-2}^* = \alpha y_{n-3} + (1-\alpha)y_{n-3}^*$ otrzymamy

$$\begin{aligned} y_n^* &= \alpha y_{n-1} + \alpha(1-\alpha)y_{n-2} + (1-\alpha)^2(\alpha y_{n-3} + (1-\alpha)y_{n-3}^*) = \\ &= \alpha y_{n-1} + \alpha(1-\alpha)y_{n-2} + \alpha(1-\alpha)^2 y_{n-3} + (1-\alpha)^3 y_{n-3}^* \end{aligned}$$

ostatecznie

$$y_n^* = \alpha y_{n-1} + \alpha(1-\alpha)y_{n-2} + \dots + \alpha(1-\alpha)^k y_{n-k+1} + \dots$$

Wagi przy poszczególnych elementach szeregu czasowego

$$\alpha > \alpha(1-\alpha) > \dots > \alpha(1-\alpha)^k > \dots$$

stanowią kolejne wyrazy ciągu geometrycznego o ilorazie $0 < 1-\alpha < 1$. Dla dużych n ich suma jest prawie równa 1 bowiem

$$\alpha + \alpha(1-\alpha) + \dots + \alpha(1-\alpha)^k + \dots = \frac{\alpha}{1-(1-\alpha)} = 1$$

Uwaga.

- a) jeśli wygładzenie szeregu czasowego (zwłaszcza dla dużych α) nie jest zadowalające to możemy powyższe wygładzanie powtórzyć,
- b) chociaż dla małych α wygładzenie jest lepsze, to nie zawsze wtedy jest najmniejszy błąd średniokwadratowy dla prognoz przeszłych, widać to w następującym przykładzie

Przykład, .

Liczba sprzedanych żarówek (tys. szt.) w hurtowni „LUMEN” w kolejnych kwartałach lat 1998-2000:

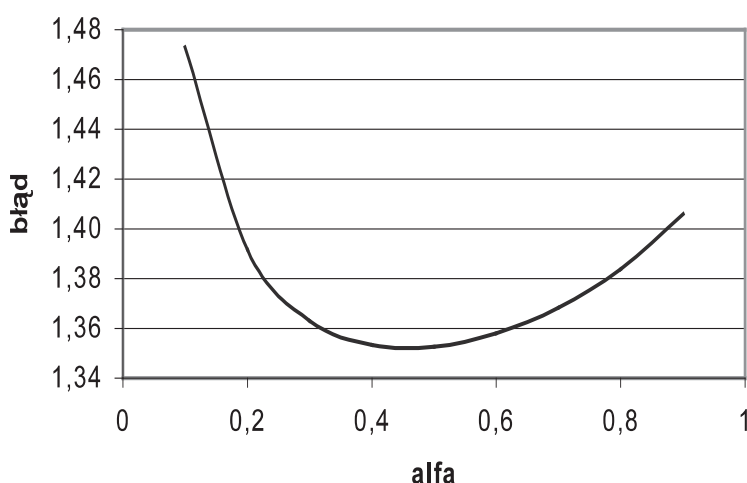
37, 36, 34, 33, 34, 33, 35, 34, 35, 33, 34, 36

Badając wielkość błędu średniokwadratowego dla różnych wartości α otrzymamy:

α	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
błąd	1,47	1,39	1,36	1,35	1,35	1,36	1,37	1,38	1,41

Jak widać najlepsze (z tego punktu widzenia) wartości α są w przedziale $0,4 \div 0,5$.

błąd średniokwadratowy (zależność od α)



Szereg czasowy z trendem. Model Holta.

Stosujemy dwa równania rekurencyjne:

I - służy do wyznaczania wygładzonych wartości szeregu czasowego w chwili t

$$F_t = \alpha y_t + (1 - \alpha)(F_{t-1} + S_{t-1})$$

II - służy do wyznaczania wygładzonych wartości przyrostu trendu w chwili t

$$S_t = \beta(F_t - F_{t-1}) + (1 - \beta)S_{t-1}$$

$\alpha, \beta \in \langle 0, 1 \rangle$ - parametry wygładzania.

Ich wartość dobieramy np. na podstawie kryterium najmniejszego błędu średniego prognoz
wygasłych s^* tzn. $\min_{\alpha, \beta} s^*(\alpha, \beta)$ gdzie

$$s^* = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - y_t^*(\alpha, \beta))^2}$$

Prognozy wygasłe obliczamy wg wzoru

$$y_{t+1}^* = F_t + S_t$$

Prognozę zmiennej Y na okres T ($T > n$) $T = n+1, n+2$ itd.

$$y_T^* = F_n + (T - n)S_n$$

w szczególności dla $T = n + 1$ mamy:

$$y_{n+1}^* = F_n + S_n$$

(wszystkie kolejne prognozy leżą na prostej $y = F_n + x \cdot S_n$)

Uwaga.

Ponieważ $y_t^* = F_{t-1} + S_{t-1}$ to

$$F_t = y_t^* + \alpha(y_t - y_t^*) \quad S_t = S_{t-1} + \alpha\beta(y_t - y_t^*)$$

Wartości początkowe F_1 i S_1 wyznaczamy wg propozycji

Propozycja	F_1	S_1
1	y_1	0
2	y_1	$y_2 - y_1$
3	Wyraz wolny liniowej funkcji trendu oszacowany na podstawie np. kilku pierwszych obserwacji	Współczynnik kierunkowy liniowej funkcji trendu oszacowany na podstawie np. kilku pierwszych obserwacji

SYMULACJE (analiza wariantowa)

Przykład

Rozpatrzmy model potęgowy popytu oszacowany na podstawie danych z lat 1990-2001

$$\hat{Y} = 100 X_1^{0,5} X_2^{-1,5}$$

Y - popyt (tys. szt/rok)

X₁ - dochody osób (zł/rok)

X₂ - cena (zł/szt)

Interpretacja parametrów.

Przy ustalonej cenie wzrost dochodów o 1% powoduje wzrost popytu o 0,5%.

Przy ustalonych dochodach wzrost ceny o 1% powoduje spadek popytu o 1,5%.

Problemy.

- a) Jaka powinna być wielkość produkcji w 2002 roku aby nie było zapasów wiedząc, że w 2002 roku dochody mogą być w granicach 10000 - 20000 a cena 6 - 8?
- b) Jaka powinna być cena w 2002 roku aby przy dochodach w granicach 10000 - 20000 sprzedać 600 000 szt?

Ad a)

Zakładamy, że parametry modelu nie zmieniają się. Ponieważ funkcja modelu jest monotoniczna ze względu na poszczególne zmienne to do określenia granic wielkości produkcji (popytu) wystarczy rozpatrzeć dolne i górne granice zakresu zmiennych objaśniających.

Dla X₁ = 10000, X₂ = 6 mamy Y= 680

Dla X₁ = 10000, X₂ = 8 mamy Y= 442

Dla X₁ = 20000, X₂ = 6 mamy Y= 962

Dla X₁ = 20000, X₂ = 8 mamy Y= 625

Zatem należy produkować co najmniej 442 000 szt. (wariant pesymistyczny), lub co najwyżej 962 000 szt. (wariant optymistyczny).

Ad b)

Obliczamy z równania modelu zmienną X_2

$$X_2 = \left(\frac{100 X_1^{0,5}}{Y} \right)^{\frac{2}{3}}$$

Dla $X_1 = 10000$, $Y = 600$ mamy $X_2 = 6,52$

Dla $X_1 = 20000$, $Y = 600$ mamy $X_2 = 8,22$

Zatem bezpieczna cena to 6,52 zł/szt. (wariant pesymistyczny).

Uwaga.

Jeśli X_1 , X_2 potraktujemy jako zmienne losowe o określonym rozkładzie to przez generowanie losowych wartości tych zmiennych można określić losową wartość popytu.

Zadanie.

Rozpatrzmy model potęgowy popytu (jak wyżej) oszacowany na podstawie danych z lat 1990-2001

$$\hat{Y} = 20 X_1^{0,7} X_2^{-0,8}$$

- a) Podać interpretację parametrów,
- b) Jaka powinna być wielkość produkcji w 2002 roku aby nie było zapasów wiedząc, że w 2002 roku dochody mogą być w granicach 20000 - 25000 a cena 3 - 4?
- c) Jaka powinna być cena w 2002 roku aby przy dochodach w granicach 10000 - 20000 sprzedać 8 000 000 szt?

TEORIA GIER

Teoria gier to teoria podejmowania decyzji w interaktywnych sytuacjach konfliktu lub współpracy. Początkowo gry utożsamiano z grami towarzyskimi lub hazardowymi, ich problematyką zajmowało się wielu matematyków i filozofów. Teoria gier jako dziedzina matematyki stosowanej zaistniała w 1944 roku dzięki monografii Johna von Neumanna i Oskara Morgensterna "Theory of Games and Economic Behavior" i szybko znalazła zastosowania w ekonomii, socjologii, psychologii, biologii, informatyce, systemach łączności, naukach politycznych, wojskowości i ekologii. Ponowne ożywione zainteresowanie teorią gier nastąpiło po roku 1994 gdy trzech matematyków otrzymało nagrodę Nobla z ekonomii za prace w tej dziedzinie (John Nash, John Harsanyi i Reinhard Selten).

Wprowadzenie, podstawowe pojęcia.

Przyjmuje się, że w grze powinny uczestniczyć przynajmniej dwie osoby (państwa, firmy, gatunki itd., uczestnikiem gry może też być natura). Każdy uczestnik (gracz) ma pewną liczbę możliwych sposobów rozegrania gry (strategii) . Wynik gry jest określony przez układ strategii wybranych przez poszczególnych graczy. Każdemu wynikowi gry odpowiada zestaw wypłat dla poszczególnych graczy, określają one wynik gry .

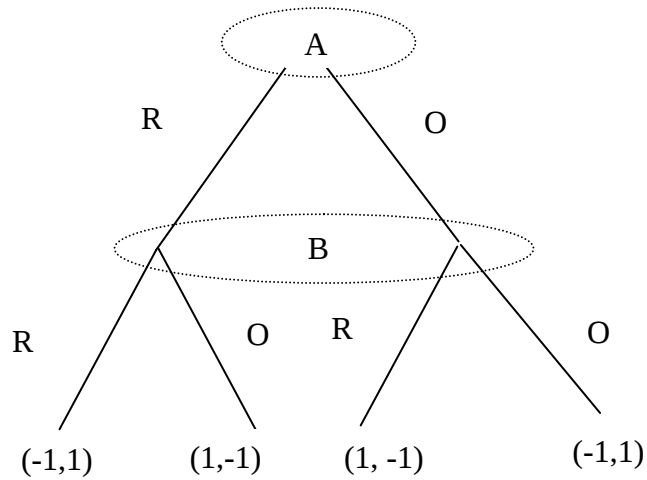
Rozpatrzmy konkretne sytuacje, które mogą być analizowane przez teorię gier.

Przykład 1

Gra w orła (O) i reszkę (R).

Gracz A wybiera orła lub reszkę. Gracz B nie znając wyboru gracza A wybiera również orła lub reszkę. Jeśli obydwoj wybrali to samo, gracz B wygrywa 1 zł od gracza A, w przeciwnym przypadku gracz A wygrywa 1 zł od gracza B.

Rozpatrywaną grę można zapisać w postaci drzewa (tzw. postać ekstensywna lub rozwinięta)



Liczby przy wierzchołkach końcowych oznaczają wartości wypłaty poszczególnych graczy (pierwsza liczba to wypłata gracza A, druga - gracza B).

Drugi sposób zapisu gry to postać macierzowa (postać normalna).

		B	
		O	R
A	O	(-1, 1)	(1, -1)
	R	(1, -1)	(-1, 1)

lub krócej

$$\begin{bmatrix} (-1, 1) & (1, -1) \\ (1, -1) & (-1, 1) \end{bmatrix} \quad (\text{tzw. zapis dwumacierzowy}).$$

Ponieważ suma wypłat w każdym przypadku wynosi zero (gra o sumie zerowej) to zapis macierzowy tej gry może być jeszcze prostszy (uwzględniamy tylko wypłaty gracza A, wypłaty gracza B są domyślnie liczbami przeciwnymi)

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Przykład 2

Dylemat więźnia.

Dwaj osobnicy podejrzani o popełnienie poważnego przestępstwa zostali zatrzymani na podstawie poszlakowych informacji i umieszczeni w oddzielnych pomieszczeniach. Otrzymują propozycje współpracy z prowadzącym śledztwo. Zatrzymany, który obciąży drugiego z zatrzymanych będzie wolny, natomiast obciążony otrzyma wyrok 8 lat więzienia. Jeżeli obaj przyznają się do winy to dostaną po 6 lat, gdy żaden się nie przyzna to zostaną skazani tylko za udział w innym drobnym przestępstwie na kary po roku więzienia. Każdy z zatrzymanych chce otrzymać jak najniższy wyrok i ma dwie strategie "przyznać się", "nie przyznać się". Rozpatrywane zagadnienie można zapisać w postaci macierzowej

$$\begin{bmatrix} (-6, -6) & (-8, 0) \\ (0, -8) & (-1, -1) \end{bmatrix}$$

Przykład 3

Zmowa kartelowa.

Dwaj producenci wytwarzają ten sam towar. Jeśli zawrą porozumienie dotyczące ceny i wielkości produkcji aby osiągnąć zysk maksymalny to łączny zysk w wysokości 200 tys zł podzielą po równo. Zwiększenie produkcji grozi obniżką ceny (zatem i zysku), ale firmy wiedzą, że jeśli tylko jedna firma zwiększy produkcję to jej zysk wyniesie 150 tys zł a rywala, który nie zmienił wielkości produkcji tylko 40 tys zł. Jeśli jednak obie firmy zwiększą produkcję to otrzymają tylko po 50 tys zł.

Każda z firm ma dwie strategie "wykonać ustaloną produkcję", "zwiększyć produkcję".

. Rozpatrywane zagadnienie jest podobne do "dylematu więźnia" i można go zapisać w postaci macierzowej

$$\begin{bmatrix} (100, 100) & (40, 150) \\ (150, 40) & (50, 50) \end{bmatrix}$$

Inne sytuacje prowadzące do gier typu "dylemat więźnia" to np.

a) dwa konkurujące ze sobą sklepy i strategie "obniżyć ceny", "nie obniżać cen ", jako wypłaty rozpatrujemy zyski poszczególnych sklepów.

b) dwa konkurujące ze sobą państwa zaangażowane w wyścig zbrojeń i strategię "zbroić się", "nie zbroić się", jako wypłaty rozpatrujemy zyski poszczególnych państw uzyskane w wyniku wygrania konfliktu.

Ogólnie przyjmuje się, że gra jest typu "dylemat więźnia" gdy jej macierz ma postać

$$\begin{bmatrix} (a, a) & (b, c) \\ (c, b) & (d, d) \end{bmatrix}$$

gdzie $c > d > a > b$, oraz niekiedy dodatkowo $d \geq (c + b) / 2$.

Zagadnienie porównywania i racjonalnego wyboru strategii rozpatrzmy analizując następującą grę rynkową.

Przykład 4

Rozpatrzmy przykład rywalizacji dwóch producentów A i B wytwarzających ten sam towar. Każda z firm ma dwie strategię a_1, b_1 - "produkcja wysoka", a_2, b_2 - "produkcja niska". Każda z firm chce osiągnąć jak największy zysk (wypłata). Poziom zysku w zależności od przyjętej strategii wyraża macierz tej gry.

$$\begin{array}{cc} & b_1 & b_2 \\ a_1 & (300, 300) & (300, 200) \\ a_2 & (200, 600) & (500, 500) \end{array}$$

Pierwsza firma nie ma **strategii dominującej**, tj. takiej, która niezależnie jak zachowa się druga firma pozwala osiągnąć najwyższy zysk. Jeśli bowiem firma B wybierze strategię b_1 to firma A powinna wybrać strategię a_1 (zysk 300 jest większy od zysku 200), jeśli natomiast firma B wybierze strategię b_2 to firma A powinna wybrać strategię a_2 (zysk 500 jest większy od zysku 300). Inaczej jest w przypadku drugiej firmy, ma ona strategię dominującą b_1 bowiem bez względu czy firma A wybierze strategię a_1 czy a_2 to firma B ma większy zysk w przypadku gdy zastosuje strategię b_1 .

Firma zachowuje się racjonalnie gdy maksymalizuje swój zysk znając zachowanie konkurencyjnej firmy. Taki wybór strategii nazywamy **równowagą**. W naszym przykładzie równowaga zachodzi przy wyborze strategii a_1 b_1 . Zauważmy, że podana para strategii jest jedyną równowagą tej gry.

Jak widać w przypadku ogólnym do prawidłowego określenia gry, w której uczestniczą dwa podmioty (dwie firmy A i B, lub dwie osoby np. Ania i Bartek, itd.) dwuosobowej (gra w postaci **normalnej**) potrzebujemy czterech elementów:

- Zbioru strategii gracza A,
- Zbioru strategii gracza B,
- Funkcji wypłaty gracza A,
- Funkcji wypłaty gracza B, oznaczmy

Jeżeli gracz A ma m strategii, gracz B - n strategii to jak widzieliśmy w powyższych przykładach funkcję wypłaty można zapisać w postaci macierzy podwójnej (ewentualnie jako parę macierzy) o wymiarach $m \times n$, na przecięciu i - tego wiersza oraz j - tej kolumny wpisujemy parę wypłat (f_A, f_B) . Zamiast macierzy podwójnej ewentualnie stosuje się zapis w postaci pary macierzy.

W teorii gier tradycyjnie przyjmuje się następującą terminologię

- Podmioty uczestniczące w grze nazywa się **graczami**, a ich rozwiązania (decyzje) - **strategiami**,
- Operacja która pozwala graczowi wybrać właściwą strategię nazywa się **regulą gry**.

Gracze klasyfikują strategie za pomocą funkcji oceniających je. Funkcje te w zależności od punktu widzenia noszą nazwy:

- użyteczności,
- wypłaty,
- straty,
- wydatków (kosztów),
- przychodów (zysków).

Będziemy stosować najczęściej funkcję **wypłaty** lub **straty**.

Rozpatrując możliwe strategie gracz stara się **minimalizować swoją stratę** czyli **maksymalizować swoją wypłatę**.

W niektórych przypadkach gracz może nie być zdecydowany, którą strategię wybrać i uzależnić wybór strategii od wyniku zdarzenia losowego (np. rzut moneta, kością itp.). Tak wybrane strategie nazywane są **strategiami mieszanymi** w odróżnieniu od zwykłych strategii zwanych **czystymi**. Zatem strategię mieszaną gracza A będziemy utożsamiali z rozkładem prawdopodobieństwa na zbiorze jego strategii czystych, tj.

S_A	a_1	...	a_m
p	p_1	...	p_m

lub krócej

$$p = (p_1, \dots, p_m)$$

gdzie

$$p_1, \dots, p_m \geq 0; \quad p_1 + \dots + p_m = 1$$

W Przykładzie 4 strategia mieszana $(1/4; 3/4)$ oznacza, że gracz A użyje strategii a_1 z prawdopodobieństwem $1/4$ a strategii a_2 z prawdopodobieństwem $3/4$; do losowego wyboru strategii może posłużyć się czworościenną kostką, urną z trzema kulami białymi i jedną kulą czarną lub generatorem liczb losowych.

Przyjrzyjmy się teraz bliżej grom o sumie zerowej (wtedy $f_A = -f_B$), zapis takiej gry upraszcza się, macierz wypłaty jest macierzą pojedynczą, jest to macierz wypłaty gracza A.

W takich grach równowagę Nasha wyznacza się prościej, okazuje się, że jest ona równoważna istnieniu **punktu siodłowego**.

Rozpatrzmy następujący przykład.

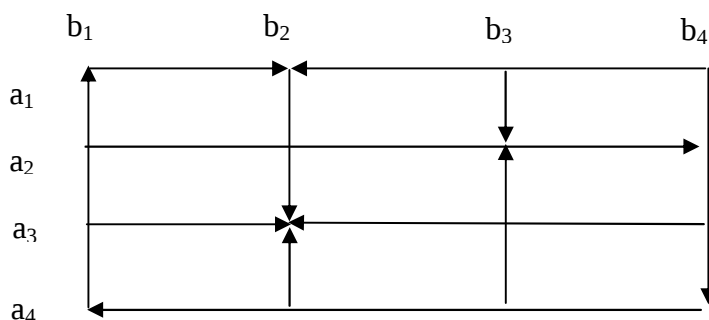
Przykład 5

Gra o macierzy wypłat

$$\begin{bmatrix} 24 & -2 & 2 & 0 \\ 10 & 2 & 14 & -40 \\ 6 & 4 & 8 & 6 \\ -32 & 0 & 0 & 32 \end{bmatrix}$$

ma równowagę Nasha, jest to para strategii (a_3, b_2) , daje ona graczowi A wypłatę 4, a graczowi B wypłatę -4.

Szukanie punktu równowagi może ułatwić następujący diagram na macierzy wypłat (w wierszach strzałki wskazują na element najmniejszy, w kolumnach na element największy).



Do podobnego wyboru strategii mogą gracze dojść w drodze następującego rozumowania.

a) Gracz A wie, że gracz B chce zminimalizować swoją stratę, wobec tego przewidując jego wybór sam wybiera w każdym wierszu liczbę najmniejszą (w tym przykładzie są to liczby -2, -40, 4, -32).

Z drugiej strony gracz A chce maksymalizować swoją wypłatę, wobec tego wybiera największą liczbę spośród poprzednio wybranych (w tym przykładzie jest to liczba 4).

b) Gracz B wie, że gracz A chce maksymalizować swoją wypłatę, wobec tego przewidując jego wybór sam wybiera w każdym wierszu liczbę największą (w tym przykładzie są to liczby 24, 4, 14, 32).

Lecz gracz B chce minimalizować swoją stratę, wobec tego wybiera najmniejszą liczbę spośród poprzednio wybranych (w tym przykładzie jest to liczba 4).

$$\begin{array}{ccccccc}
 & & & & & \min & \\
 & & & & & -2 & \\
 \left[\begin{array}{cccc} 24 & -2 & 2 & 0 \\ 10 & 2 & 14 & -40 \\ 6 & 4 & 8 & 6 \\ -32 & 0 & 0 & 32 \end{array} \right] & & -40 & & & & \\
 & & [4] & \max & & & \\
 & & -32 & & & & \\
 \max & 24 & [4] & 14 & 32 & & \\
 & & \min & & & &
 \end{array}$$

Jeśli rozumowanie a) i b) prowadzi do wyboru tych samych strategii to wyznaczona wartość macierzy wypłaty nazywamy punktem siodłowym. Tak wybrany punkt wskazuje na równowagę w tej grze.

Zatem punkt siodłowy spełnia warunek

$$\min_j \max_i a_{ij} = \max_j \min_i a_{ij}$$

Dostrzeżony w tym przykładzie związek między równowagą Nasha i punktem siodłowym jest powszechny bowiem zachodzi następujące twierdzenie dla gier o sumie zerowej..

Twierdzenie

Gra o macierzy wypłat $[a_{ij}]_{m \times n}$ ma równowagę Nasha w strategiach czystych wtedy i tylko wtedy gdy istnieje punkt siodłowy tej gry.

Są gry nie posiadające punktu siodłowego oraz gry w których jest więcej niż jeden punkt siodłowy, świadczą o tym następujące przykłady gier.

Przykład 6

a) gra o macierzy wypłat

$$\left[\begin{array}{cccc} 8 & [4] & 10 & [4] \\ 4 & 2 & -2 & -40 \\ 6 & [4] & 8 & [4] \\ -32 & 0 & 32 & 2 \end{array} \right]$$

ma cztery punkty siodłowe (wyróżnione elementy macierzy)

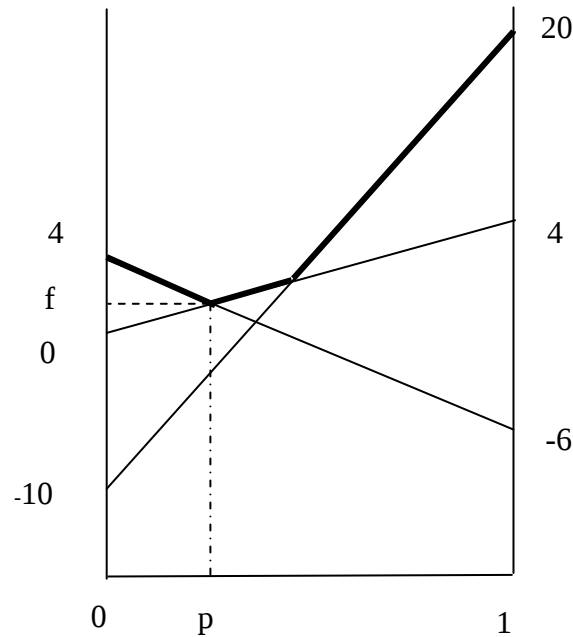
b) gra o macierzy wypłat

$$\left[\begin{array}{cc} 4 & -6 \\ 0 & 4 \\ -10 & 20 \end{array} \right]$$

nie ma punktu siodłowego.

Jeżeli jeden z graczy ma tylko dwie strategie to możemy poszukiwać punktu siodłowego w zbiorze strategii mieszanych metodą graficzną w następujący sposób (omówimy go dla Przykładu 6):

Na dwóch równoległych pionowych osiach oddalonych od siebie o jednostkę odkładamy odpowiednio punkty z pierwszej i drugiej kolumny macierzy wypłat. Następnie łączymy odcinkami punkty z tych samych wierszy.



Wybieramy najniższy punkt na górnej łamanej (strategia minmax) i rzutujemy go na odcinek $[0, 1]$. Rzut tego samego odcinka na oś rzędnych określa wartość gry f . Strategia mieszana dla gracza B będzie wyznaczona przez $(p, 1-p) = (2/7, 5/7)$. Wartość gry $f = 8/7$. Strategia mieszana dla A ma postać $(2/7, 5/7, 0)$.

DANE STATYSTYCZNE

DANE-1

Member States' contribution to the Intra-EU27 trade

Wkład państw członkowskich

In 1000 million of ECU/EURO

	2000	2001	2002	2003	2004	2005	2006	2007	2008
EU-27-total	1805.8	1872.8	1897.4	1914.5	2071.8	2215.0	2497.3	2647.8	2701.7
Belgium	156.6	165.6	172.4	174.4	190.1	206.2	224.1	240.4	249.4
Bulgaria	3.0	3.5	3.8	4.2	5.0	5.5	7.1	8.2	9.2
CzechRepublic	27.1	32.2	34.9	37.6	48.3	53.7	64.8	76.2	84.4
Denmark	39.2	40.1	42.4	41.3	43.8	48.4	52.4	52.6	55.6
Germany	386.6	406.0	412.7	431.1	472.3	501.6	561.4	623.9	633.0
Estonia	3.0	3.0	3.0	3.3	3.8	4.9	5.1	5.6	5.9
Ireland	54.3	59.4	61.6	51.2	53.0	56.2	54.8	56.3	53.7
Greece	7.9	8.1	6.7	7.7	7.9	8.6	10.6	11.2	11.0
Spain	91.1	96.9	99.4	103.9	109.2	112.1	121.1	130.8	124.4
France	230.0	231.9	228.4	231.1	239.8	236.5	258.7	261.0	259.5
Italy	160.2	166.6	163.9	165.0	176.0	183.7	203.1	215.4	213.9
Cyprus	0.3	0.3	0.3	0.3	0.5	0.9	0.8	0.7	0.8
Latvia	1.6	1.8	1.9	2.0	2.5	3.2	3.6	4.4	4.7
Lithuania	2.9	3.5	3.8	3.9	5.0	6.2	7.2	8.1	9.7
Luxembourg	7.9	9.6	9.5	10.5	11.8	13.6	16.3	14.5	15.4
Hungary	25.5	28.5	30.8	32.1	37.1	40.9	47.5	55.0	57.1
Malta	0.9	1.1	1.0	1.0	1.0	1.0	1.1	1.1	0.9
Netherlands	205.3	210.0	207.5	210.4	229.5	260.7	292.3	313.8	338.7
Austria	54.8	59.2	62.3	64.7	70.0	72.3	78.4	86.7	88.8
Poland	27.9	32.6	35.3	38.9	48.5	56.5	69.7	80.7	88.5
Portugal	21.5	21.9	22.3	22.8	23.0	24.5	26.7	28.8	28.0
Romania	8.1	9.6	10.8	11.8	14.1	15.6	18.2	21.3	23.7
Slovenia	6.9	7.3	7.5	7.7	8.9	10.5	12.7	15.2	15.8
Slovakia	11.5	12.7	13.6	16.6	19.3	22.4	29.0	36.8	41.2
Finland	31.5	29.2	29.2	28.3	28.7	29.9	35.2	37.3	36.6
Sweden	56.9	49.8	50.5	53.0	58.5	62.1	70.8	75.4	74.9
UKingdom	183.5	182.4	181.9	160.0	164.2	177.4	224.9	186.4	177.3

DANE-2

Total general government expenditure

millions of euro

Suma wydatków rządowych.

	2000	2001	2002	2003	2004	2005	2006	2007	2008
EU27	4,161,480	4,425,685	4,640,864	4,777,467	4,968,727	5,184,519	5,406,515	5,649,660	5,849,801
Belgium	123,784	127,455	133,567	140,786	143,385	157,731	154,235	161,882	171,981
Bulgaria	5,831	6,139	6,697	7,160	7,891	8,606	9,217	11,979	12,775
CzechRepublic	25,715	30,717	37,047	38,292	39,838	45,070	49,737	54,180	62,921
Denmark	93,002	97,084	100,839	103,793	107,531	109,507	112,759	115,488	120,740
Germany	930,400	1,005,060	1,030,840	1,049,210	1,041,210	1,050,450	1,052,290	1,070,090	1,094,380
Estonia	2,225	2,426	2,781	3,035	3,290	3,775	4,481	5,418	6,485
Ireland	32,983	39,023	43,773	46,555	50,215	54,644	60,298	68,128	76,084
Greece	64,406	66,318	70,614	77,143	84,317	85,593	90,020	100,353	109,160
Spain	246,542	262,982	283,597	300,643	327,015	349,383	377,876	407,849	443,172
France	744,253	772,060	815,144	849,587	883,073	921,454	952,121	991,196	1,027,695
Italy	550,032	599,587	613,734	645,251	664,303	688,306	723,388	739,945	766,134
Cyprus	3,728	4,121	4,496	5,305	5,445	5,952	6,375	6,724	7,454
Latvia	3,172	3,222	3,533	3,471	3,998	4,630	6,127	7,586	9,120
Lithuania	4,835	4,996	5,228	5,472	6,052	6,958	8,053	9,933	12,026
Luxembourg	8,270	8,607	9,964	10,794	11,684	12,570	13,083	13,550	14,920
Hungary	24,203	28,144	36,284	36,650	40,248	44,420	46,718	50,279	52,448
Malta	1,730	1,851	1,937	2,112	2,053	2,143	2,235	2,321	2,565
Netherlands	184,612	203,063	214,960	224,621	226,403	229,965	246,356	256,918	270,262
Austria	108,175	109,648	111,512	115,011	125,720	122,004	127,194	131,824	137,308
Poland	76,293	92,913	92,703	85,561	87,053	106,107	119,194	130,815	156,011
Portugal	52,740	57,432	59,945	63,057	67,040	71,009	71,944	74,697	76,335
Romania	15,638	16,324	17,020	17,589	20,460	26,698	34,489	45,296	52,773
Slovenia	10,016	10,800	11,365	11,938	12,439	12,998	13,823	14,627	16,189
Slovakia	11,213	10,481	11,703	11,835	12,807	14,692	16,447	18,861	22,668
Finland	63,871	66,750	70,222	72,995	76,267	78,986	81,343	84,980	90,083
Sweden	148,132	139,508	149,923	156,998	159,849	162,694	169,604	174,055	174,504
UnitedKingdom	625,679	658,975	701,438	692,604	759,143	808,176	857,112	900,688	863,608
Iceland	3,944	3,761	4,193	4,429	4,696	5,539	5,545	6,312	5,882
Norway	77,232	84,230	96,029	95,939	94,616	102,310	108,600	116,297	123,775

DANE-3

Long term interest rates, annual averages

Długoterminowe stopy procentowe (średnia roczna)

	2000	2001	2002	2003	2004	2005	2006	2007	2008
EuropeanUnion	5.44	5.01	4.93	4.24	4.39	3.70	4.03	4.57	4.55
Belgium	5.59	5.13	4.99	4.18	4.15	3.43	3.82	4.33	4.42
Bulgaria	:	:	:	6.45	5.36	3.87	4.18	4.54	5.38
CzechRepublic	:	6.31	4.88	4.12	4.82	3.54	3.80	4.30	4.63
Denmark	5.64	5.08	5.06	4.31	4.30	3.40	3.81	4.29	4.28
Germany	5.26	4.80	4.78	4.07	4.04	3.35	3.76	4.22	3.98
Estonia	10.48	10.15	8.42	5.25	4.39	4.17	5.01	6.09	8.16
Ireland	5.51	5.01	5.01	4.13	4.08	3.33	3.77	4.31	4.53
Greece	6.10	5.30	5.12	4.27	4.26	3.59	4.07	4.50	4.80
Spain	5.53	5.12	4.96	4.12	4.10	3.39	3.78	4.31	4.37
France	5.39	4.94	4.86	4.13	4.10	3.41	3.80	4.30	4.23
Italy	5.58	5.19	5.03	4.25	4.26	3.56	4.05	4.49	4.68
Cyprus	:	7.63	5.70	4.74	5.80	5.16	4.13	4.48	4.60
Latvia	:	7.57	5.41	4.90	4.86	3.88	4.13	5.28	6.43
Lithuania	:	8.15	6.06	5.32	4.50	3.70	4.08	4.55	5.61
Luxembourg	5.52	4.86	4.70	4.03	4.18	3.37	3.92	4.56	4.61
Hungary	:	7.95	7.09	6.82	8.19	6.60	7.12	6.74	8.24
Malta	:	6.19	5.82	5.04	4.69	4.56	4.32	4.72	4.81
Netherlands	5.40	4.96	4.89	4.12	4.10	3.37	3.78	4.29	4.23
Austria	5.56	5.07	4.97	4.15	4.15	3.39	3.80	4.29	4.26
Poland	:	10.68	7.36	5.78	6.90	5.22	5.23	5.48	6.07
Portugal	5.59	5.16	5.01	4.18	4.14	3.44	3.91	4.42	4.52
Romania	:	:	:	:	:	:	7.23	7.13	7.70
Slovenia	:	:	8.72	6.40	4.68	3.81	3.85	4.53	4.61
Slovakia	:	8.04	6.94	4.99	5.03	3.52	4.41	4.49	4.72
Finland	5.48	5.04	4.98	4.13	4.11	3.35	3.78	4.29	4.29
Sweden	5.37	5.11	5.30	4.64	4.43	3.38	3.71	4.17	3.89
UnitedKingdom	5.33	5.01	4.91	4.58	4.93	4.46	4.37	5.06	4.50
United States	6.03	5.01	4.60	4.00	4.26	4.28	4.79	4.63	3.65
Japan	1.76	1.34	1.27	0.99	1.50	1.39	1.74	1.68	1.49

DANE-4

Euro exchange rates, annual averages (1 € = ... National currency)

Kursy walut.

			2000	2001	2002	2003	2004	2005	2006	2007	2008
czk	Czech	Koruna	35.599	34.068	30.804	31.846	31.891	29.782	28.342	27.766	24.946
dkk	Danish	Krone	7.454	7.452	7.431	7.431	7.440	7.452	7.459	7.451	7.456
EEK	Estonian	Kroon	15.647	15.647	15.647	15.647	15.647	15.647	15.647	15.647	15.647
lvi	Latvian	Lats	0.559	0.560	0.581	0.641	0.665	0.696	0.696	0.700	0.703
ltl	Lithuanian	Litas	3.695	3.582	3.459	3.453	3.453	3.453	3.453	3.453	3.453
huf	Hungarian	forint	260.04	256.59	242.96	253.62	251.66	248.05	264.26	251.35	251.51
pln	Polish	Zloty	4.008	3.672	3.857	4.400	4.527	4.023	3.896	3.784	3.512
skk	Slovak	Koruna	42,602	43,3	42,694	41,489	40,022	38,599	37,234	33,775	31,262
sek	Swedish	Krona	8.445	9.255	9.161	9.124	9.124	9.282	9.254	9.250	9.615
gbp	UK	PoundSt	0.609	0.622	0.629	0.692	0.679	0.684	0.682	0.684	0.796
isk	Iceland	Krona	72.58	87.42	86.18	86.65	87.14	78.23	87.76	87.63	143.83
nok	Norwegian	Krone	8.113	8.048	7.509	8.003	8.370	8.009	8.047	8.017	8.224
chf	Swiss	Franc	1.558	1.511	1.467	1.521	1.544	1.548	1.573	1.643	1.587
bgn	Bulgarian	Lev	1.952	1.948	1.949	1.949	1.953	1.956	1.956	1.956	1.956
ron	Romanian	leu	1.992	2.600	3.127	3.755	4.051	3.621	3.526	3.333	3.678
jpy	Japan	Yen	99.47	108.68	118.06	130.97	134.44	136.85	146.02	161.25	152.45
usd	UnitedStates	Dollar	0.924	0.896	0.946	1.131	1.244	1.244	1.256	1.371	1.471

DANE-5

Employment levels (thousand persons)

Poziom zatrudnienia

		2001	2002	2003	2004	2005	2006	2007	2008
eu27	EuropeanUnion	211,460	212,239	213,022	214,489	216,557	220,102	224,143	226,208
ea15	Euroarea	135,770	136,753	137,338	138,452	139,779	142,048	144,653	145,768
be	Belgium	4,150	4,144	4,146	4,175	4,229	4,288	4,365	4,436
bg	Bulgaria	3,215	3,222	3,317	3,403	3,495	3,612	3,714	3,836
cz	CzechRepublic	4,963	4,991	4,923	4,940	4,992	5,072	5,207	5,268
dk	Denmark	2,785	2,786	2,756	2,739	2,767	2,822	2,898	2,928
de	Germany	39,316	39,096	38,726	38,880	38,851	39,097	39,768	40,330
ee	Estonia	577	584	593	593	604	637	642	643
ie	Ireland	1,748	1,779	1,814	1,870	1,958	2,042	2,115	2,096
gr	Greece	4,261	4,356	4,401	4,503	4,546	4,642	4,702	4,759
es	Spain	16,931	17,338	17,878	18,510	19,267	20,024	20,626	20,532
fr	France	24,765	24,919	24,950	24,977	25,116	25,362	25,705	25,841
it	Italy	23,393	23,793	24,150	24,256	24,396	24,874	25,184	25,263
cy	Cyprus	322	328	341	354	366	373	385	395
lv	Latvia	965	980	997	1,008	1,026	1,074	1,112	1,120
lt	Lithuania	1,346	1,395	1,426	1,425	1,461	1,487	1,529	1,522
lu	Luxembourg	278	287	293	299	308	319	333	349
hu	Hungary	3,855	3,856	3,906	3,880	3,880	3,905	3,900	3,855
mt	Malta	149	150	151	150	152	154	159	163
nl	Netherlands	8,282	8,324	8,283	8,211	8,252	8,404	8,613	8,743
at	Austria	3,816	3,812	3,809	3,862	3,919	3,974	4,046	4,117
pl	Poland	15,242	14,960	13,606	13,773	14,075	14,530	15,174	15,783
pt	Portugal	5,121	5,151	5,122	5,117	5,100	5,126	5,125	5,147
ro	Romania	:	9,574	9,569	9,410	9,267	9,331	9,372	9,396
si	Slovenia	909	923	919	922	921	935	963	990
sk	Slovakia	2,037	2,038	2,060	2,056	2,084	2,132	2,177	2,237
fi	Finland	2,331	2,353	2,356	2,365	2,397	2,440	2,493	2,530
se	Sweden	4,391	4,393	4,368	4,337	4,349	4,423	4,518	4,559
uk	UnitedKingdom	27,711	27,922	28,189	28,489	28,779	29,030	29,226	29,439
hr	Croatia	1,465	1,526	1,535	1,561	1,573	1,564	1,618	1,635
mk	Macedonia	483	480	471	460	470	485	506	522
no	Norway	2,328	2,337	2,313	2,323	2,352	2,437	2,538	2,616
ch	Switzerland	4,155	4,173	4,167	4,178	4,196	4,291	4,347	4,377
us	UnitedStates	139,222	138,807	140,084	141,569	143,980	146,678	148,295	147,603
jp	Japan	64,761	63,747	63,539	63,676	63,918	64,198	64,500	64,228

DANE-6

Supply, transformation - oil - monthly data, Crude oil and Petroleum Products, Thousands of tons

geo/time	2009M01	2009M02	2009M03	2009M04	2009M05	2009M06	2009M07	2009M08
EUnion (27)	9451	8660	9131	9259	8884	8647	8800	7216
Belgium	-	-	-	-	-	-	-	-
Bulgaria	2	2	2	2	2	2	2	2
Czech Republic	31	31	35	30	33	33	38	38
Denmark	1201	1067	1140	1110	1081	1066	1112	1061
Germany	374	355	424	391	378	376	391	379
Estonia	-	-	-	-	-	-	-	-
Ireland	-	-	-	-	-	-	-	-
Greece	5	5	0	0	1	0	0	10
Spain	47	15	9	8	8	9	12	11
France	200	198	208	224	189	216	242	193
Italy	466	391	496	474	468	411	400	458
Cyprus	-	-	-	-	-	-	-	-
Latvia	-	-	-	-	-	-	-	-
Lithuania	14	11	14	16	15	16	17	15
Luxembourg	-	-	-	-	-	-	-	-
Hungary	116	98	143	102	104	99	96	109
Netherlands	189	190	-220	185	173	146	184	167
Austria	83	77	86	77	77	78	77	77
Poland	60	53	57	54	57	56	42	41
Portugal	-	-	-	-	-	-	-	-
Romania	399	379	375	383	328	364	379	385
Slovenia	-	-	-	-	-	-	-	-
Slovakia	10	11	12	12	12	14	12	22
Finland	-	-	-	-	-	-	-	-
Sweden	-	-	-	-	-	-	-	-
United Kingdom	6254	5778	6349	6191	5958	5760	5796	4248
Croatia	81	73	81	77	81	78	80	78
Turkey	166	160	203	198	205	200	216	215
Norway	9314	8689	9756	8929	8397	7703	9201	8099

DANE-7

Students (1 000),
Liczba studentów

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
EU (27 countries)	15013,2	15569,5	15920,8	16517,3	17139,3	17761,8	18232,9	18530,2	18782,5	18884,2
Belgium	:	351,8	355,7	359,3	367	374,7	386,1	389,5	394,4	393,7
Bulgaria	260,5	270,1	261,3	247	228,4	230,5	228,5	237,9	243,5	258,7
Czech Republic	215	231,2	253,7	260	284,5	287	318,9	336,3	337,4	362,6
Denmark	183,3	190	189,2	190,8	195,3	201,7	217,1	232,3	228,9	232,2
Germany	2097,7	2087	2054,8	2083,9	2159,7	2242,4	2330,5	2268,7	2289,5	2278,9
Estonia	43,1	48,7	53,6	57,8	60,6	63,6	65,7	67,8	68,3	68,8
Ireland	142,8	151,1	160,6	166,6	176,3	181,6	188,3	186,6	186	190,3
Greece	374,1	387,9	422,3	478,2	529,2	561,5	597	646,6	653	602,9
Spain	1746,2	1786,8	1829	1833,5	1832,8	1840,6	1839,9	1809,4	1789,3	1777,5
France	2027,4	2012,2	2015,3	2031,7	2029,2	2119,1	2160,3	2187,4	2201,2	2179,5
Italy	1869,1	1797,2	1770	1812,3	1854,2	1913,4	1986,5	2015	2029	2033,6
Cyprus	:	10,8	10,4	11,9	13,9	18,3	20,8	20,1	20,6	22,2
Latvia	70,2	82	91,2	102,8	110,5	118,9	127,7	130,7	131,1	129,5
Lithuania	96,4	107,4	121,9	135,9	148,8	167,6	182,7	195,4	198,9	199,9
Luxembourg	1,8	2,7	2,4	2,5	3	3,1	:	:	2,7	:
Hungary	254,7	279,4	307,1	330,5	354,4	390,5	422,2	436	438,7	431,6
Malta	:	5,8	6,3	7,4	7,3	8,9	7,9	9,4	8,9	9,8
Netherlands	461,4	469,9	487,6	504	516,8	526,8	543,4	565	579,6	590,1
Austria	247,5	252,9	261,2	264,7	223,7	229,8	238,5	244,4	253,1	261
Poland	1191,1	1399,1	1579,6	1775	1906,3	1983,4	2044,3	2118,1	2145,7	2146,9
Portugal	351,8	356,8	373,7	387,7	396,6	400,8	395,1	380,9	367,3	366,7
Romania	360,6	407,7	452,6	533,2	582,2	643,9	685,7	738,8	835	928,2
Slovenia	68,1	79,1	83,8	91,5	99,2	101,5	104,4	112,2	114,8	115,9
Slovakia	112,8	122,9	135,9	143,9	152,2	158,1	164,7	181,4	197,9	218
Finland	250	262,9	270,2	279,6	283,8	291,7	299,9	306	309	309,2
Sweden	280,7	335,1	346,9	358	382,9	414,7	429,6	426,7	422,6	413,7
United Kingdom	1938,4	2081	2024,1	2067,3	2240,7	2287,8	2247,4	2287,5	2336,1	2362,8

Croatia	:	:	:	:	:	121,7	126,3	134,7	136,6	140
Macedonia	32	35,1	36,9	40,2	44,7	45,6	46,6	49,4	48,4	58,2
Turkey	:	1464,7	1015,4	1607,4	1677,9	1918,5	1972,7	2106,4	2342,9	2453,7
Iceland	8,1	8,5	9,7	10,2	11,6	13,3	14,7	15,2	15,7	15,8
Liechtenstein	:	:	0,5	:	:	0,4	0,5	0,5	0,6	0,7
Norway	183	187,5	190,9	190,1	197,1	212,4	213,8	213,9	214,7	215,2
Switzerland	:	:	:	:	170,1	186	195,9	199,7	205	213,1
United States	13284	13769,4	13202,9	13595,6	15928	16611,7	16900,5	17272	17487,5	17758,9
Japan	3963,7	3940,8	3982,1	3972,5	3966,7	3984,4	4031,6	4038,3	4084,9	4032,6

DANE-8

Prison population

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Belgium	8176	7889	8688	8544	8605	9308	9249	9330	9573	9950
Bulgaria	10779	10787	9424	9283	9607	10056	10935	11399	11452	10792
Czech Republic	22085	23395	22418	20971	16597	17180	18303	19003	18904	19110
Denmark	3422	3477	3382	3236	3435	3641	3767	4041	3932	3646
Germany	69917	69214	70252	70203	75025	79183	79329	79519	77166	73319
Estonia	4791	4379	4679	4803	4775	4352	4575	4410	4310	3486
Ireland	2620	2741	2887	3025	3028	2986	3138	3022	3053	:
Greece	7129	7525	8038	8343	8284	8555	8760	9871	10280	:
Spain	44747	45384	45309	46594	50537	54497	58087	60707	64215	67100
France	53667	53948	48835	47005	53463	55407	59246	59197	59522	60403
Italy	49173	52870	54039	55751	55670	54237	56068	59523	39005	48693
Cyprus	226	247	287	369	351	355	546	536	599	671
Latvia	10070	8815	8831	8673	8358	8222	7666	6998	6636	6548
Lithuania	13813	14207	8667	10750	11345	8957	7838	7951	7982	7770
Luxembourg	392	386	400	341	391	455	577	735	738	666
Hungary	14366	15110	15539	17275	17838	16507	16543	15720	14740	14743
Malta	260	230	246	257	283	278	277	294	376	382
Netherlands	11760	11870	11760	12410	13060	13980	16455	17600	16230	14450
Austria	6891	6877	6896	6915	7511	7816	9000	8955	8780	8887
Poland	59180	54842	65336	80004	80990	80692	79344	82656	87669	90199
Portugal	14330	13086	12728	13210	13772	13635	12956	12687	12446	11587
Romania	52149	49790	48267	49840	48075	42815	39031	36700	34038	29390
Slovenia	793	935	1136	1155	1120	1099	1126	1132	1127	1336
Slovakia	6897	6904	7136	7509	7849	8829	9504	9289	8657	8235
Finland	2772	2663	2887	3110	3469	3463	3535	3883	3477	3370
Sweden	5290	5484	5678	6089	6506	6755	7332	7021	7153	6740
Scotland	6029	5975	5883	6186	6475	6621	6779	6857	7183	7376
Northern Ireland	1454	1158	1011	872	1029	1128	1219	1325	1501	1484
Croatia	:	:	2623	2679	2641	2803	3010	3485	3833	4290
Macedonia	1121	1238	1425	1398	1291	1545	1791	2081	2090	:

Turkey	60096	67676	50628	55804	59512	63796	58016	55966	70524	90732
Iceland	:	:	78	110	107	112	115	119	119	115
Liechtenstein	65	71	69	86	67	67	59	62	48	38
Norway	2466	2512	2548	2762	2832	2944	3028	3124	3250	3420
Switzerland	5648	5818	5727	5137	4937	5214	5977	6137	5888	5715
Montenegro	:	:	671	669	715	744	802	816	852	961
Serbia	:	:	:	:	:	:	7653	8078	7862	8970
Russian Federation	996913	1006085	923600	979285	860640	847004	763054	797400	:	:
South Africa	143003	158682	166587	172203	181553	184576	186467	175704	157995	164957
Canada	36971	35552	35533	35736	35868	34154	34244	:	:	:
United States	1816931	1893115	1935753	1961247	2033331	2081580	2135335	2183152	2245189	2375615
Japan	51986	53947	58747	63415	67354	71889	75289	77932	:	:
Australia	19906	21538	21714	22458	22492	23555	24171	25353	25790	25224
New Zealand	5450	5661	5720	5887	5738	6059	6556	7159	7686	8235

L.Kowalski 15.02.2010